ESSAYS ON LIFE CYCLE DYNASTIC DISCRETE CHOICE MODELS

by

Mehmet Ali Soytas

BSc., Bogazici University, Istanbul, Turkey, 2001

M.A., Marmara University, Istanbul, Turkey, 2005

Submitted to the Graduate Faculty of

the Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Mehmet Ali Soytas

It was defended on

June 29, 2011

and approved by

Jean-Francois Richard, Distinguished University Professor of Economics, University of Pittsburgh

George-Levi Gayle, Associate Professor of Economics and Strategy, Carnegie Mellon University

Limor Golan, Associate Professor of Economics, Carnegie Mellon University

Randall Walsh, Associate Professor of Economics, University of Pittsburgh

David DeJong, Professor of Economics, University of Pittsburgh

Dissertation Director: Jean-Francois Richard, Distinguished University Professor of Economics, University of Pittsburgh

George-Levi Gayle, Associate Professor of Economics and Strategy, Carnegie Mellon University

ESSAYS ON LIFE CYCLE DYNASTIC DISCRETE CHOICE MODELS

Mehmet Ali Soytas, PhD

University of Pittsburgh, 2011

Models of dynastic households have been traditionally used to analyze persistence in earnings and wealth across generations, more recently to study patterns of wealth and fertility, transfers to children and education choices. However most of those models have looked at the theoretical outcomes and there are some limited calibration studies. Some other literature follows the regression based techniques to answer empirical questions regarding the generational transfers. In Chapter 2 -co-authored with Gayle and Golan- we develop an estimator that makes the structural estimation of dynastic models feasible. We propose an estimation framework for dynastic models which allows the estimation of the problem in several steps. Our estimator compared to the full solution structural estimation known as the Nested Fixed Point Algorithm (NFXP) performs comparable in small samples while reducing the computation time considerably. A Monte Carlo exercise compares our estimator to the NFXP. We show that the alternative representation of the continuation value of the problem enables us to apply the Hotz and Miller (1993) estimation to the dynastic problem.

Using data of two generations from the PSID, Chapter 3 estimates a dynastic life-cycle model with endogenous fertility, labor supply and inter-generational transfers. This chapter uses data on time spent with children and measures outcomes in terms of education. Education and skills both affect the children's earnings and marriage market outcomes stochastically. We contribute to the literature by measuring the returns in a life-cycle dynastic model in which fertility and time spent with children is endogenous and the different aspects of returns to investment (i.e. education and skill) in children are aggregated and measured in terms of their life-time utility. We model couples decisions as a noncooperative game and

solve for a Markov Perfect Equilibrium (MPE) in pure strategies. Therefore the valuation functions of the dynastic model are not only the optimal solution to the problem given the state variables for the individual, but they are the best response valuation functions given the spouse's choice. This requires an equilibrium choice which we assume as MPE in pure strategies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my advisors Jean-François Richard and George-Levi Gayle for their guidance and inspiration throughout my graduate career. George was the person who inspired me to work in the field of empirical microeconomics and I have learned a lot working with him. It was a great pleasure and chance to have Jean-François Richard as my advisor. His help and guidance played an important role at different stages of this dissertation. Also special thanks to Limor Golan who was practically one of my advisors. I also thank the other committee members, David DeJong and Randall Walsh for helpful comments and discussions. Finally, I would like to express my deep gratitude to my mother and my lovely wife Serap for their love and support.

1.0 INTRODUCTION

Models of dynastic households have been traditionally used to analyze persistence in earnings and wealth across generations (e.g. Loury (1981)(53) Laitner (1992)(50), Barro and Becker (1988, 1989)(8)(9), and more recently to study patterns of wealth and fertility, transfers to children and education choices. A number of recent empirical papers study models of intergenerational transfers. Rios-Rull and Marcos (2002)(70) study the returns of parental investment in children's education, their earnings and marriage market, Doepke and Tertilt (2009)(29) allow the returns on investment in children's human capital to depend on the parents' education. Echevarria and Merlo (1999)(30) model household bargaining in which gender gap in parental investment in education of the children arises endogenously. Doepke (2004)(28) extends Barro and Becker (1989)(9) model by allowing uncertainty over the number of children. Albanesi and Olivetti (2010)(2) link the pattern of the baby boom and bust to the improvement in maternal health. Being theoretically useful, however structural estimation of these models face computational obstacles. The problem can be solved with a nested fixed point algorithm (NFXP), it becomes computationally intensive quickly, limiting the scope of the problem that can be analyzed.

In Chapter 2 - joint with George Gayle and Limor Golan - we propose a framework for estimation of dynastic intergenerational models by developing a new representation of the problem in terms of the model's primitives and choice probabilities which allows for the estimation of the problem in several steps. The difficulty of estimating the model is due to the non-standard nature of the problem. While the problem can be solved with a nested fixed point algorithm, it becomes computationally intensive quickly. Our intergenerational model has finite (T) periods in the lifecycle in each generation and infinitely many generations are linked by the altruistic preferences. In this respect our model is close to Laitner (1992)(50).

Laitner's framework however, does not fit into a finite horizon dynamic discrete choice model since in the last period T, there is a continuation value associated with the next generation's problem which is linked to current generation by the transfers and the discount factor. Therefore we develop a representation for the next generation's continuation value that allows us to treat the problem as a standard T period problem which can be solved by backwards induction. We show that the alternative representation of the continuation value of the intergenerational problem enables us to derive the necessary representation and apply the Hotz and Miller estimation technique for single agent problems to the dynastic problem. The framework we developed can be used to estimate a large class of dynastic models with endogenous choices in the lifecyle which have intergenerational consequences. To the best of our knowledge, this is the first paper to structurally estimated dynastic model with altruistic preferences.

The framework we developed can be used to estimate a large class of dynastic models with endogenous labor supply, transfers to children, fertility, and household bargaining. Loury (1981)(53) is one of the first to model the effect of parental income on offspring's productivity, Laitner (1992)(50) incorporates lifecycle decisions into the intergenerational framework, and Barro and Becker (1988,1989)(8)(9) analyze fertility decisions. Alverez (1999)(5) provides a general framework which incorporates fertility and transfers. The third chapter structurally estimate a dynastic models with altruistic preferences. It uses the developed estimation technique to empirically address the quantity quality trade off regarding the children. Using data of two generations from the Panel Study of Income Dynamics (PSID), chapter 3 develops and estimates a dynastic life-cycle model with endogenous fertility, labor supply and intergenerational transfers. Specifically, individuals choose fertility, labor supply sequentially in their lifecycle. The focus of the empirical applications is on the effect of parental choices and characteristics on children's labor market outcomes, and the quantity-quality trade-offs involved in fertility decisions across education groups, and households' characteristics. We estimate the utility parameters of the spouses within the households. Also we estimate the generation discount factor and find that the marginal value of children is decreasing in the number of previous birth. Our paper is related to Kang (2010)(48) which estimates a lifecycle model of parental transfers, fertility and labor supply capturing the quantity-quality trade-off, in a model without dynastic component. Our paper, however, uses data on maternal time invested in children and focuses on estimation of the intergenerational discount factor. Our model belongs to the literature of dynastic models.

Chapter 3 -joint with George Gayle and Limor Golan - adds the time investment in children of father and household bargaining to the model and estimates the effects of parental time investment on the future education outcomes of children. Individuals choose fertility, labor supply and time investment in children sequentially. Echevarria and Merlo (1999)(30) investigates gender differences in education where men and women of each generation bargain over consumption, number of children and investment in education of their children. Our model builds on the literature above and incorporates fertility, labor supply and transfers decisions, made sequentially by households in a noncooperative game theoretical framework. Chapter 3 uses data on time spent with children and measures outcomes in terms of education, which in turns affects labor market skill. Education and skills both affect the children's earnings and marriage market outcomes stochastically. We contribute to the literature by measuring the returns in a life-cycle dynastic model in which fertility and time spent with children is endogenous and the different aspects of returns to investment (i.e. education, skill and marriage market outcomes) in children are aggregated and measured in terms of their life-time utility. We model couples decisions as a noncooperative game and solve for a Markov Perfect Equilibrium (MPE) in pure strategies. We assume that each period decisions are made in two stages. In the first stage labor supply, investment, transfers to children are chosen by each individual and birth decisions by the females simultaneously. In a second stage consumption allocation is made according to the sharing rules. Therefore the valuation functions of the dynastic model are not only the optimal solution to the problem given the state variables for the individual, but they are the best response valuation functions given the spouse's choice. This requires an equilibrium choice which we assume as Markov Perfect Equilibrium (MPE) in pure strategies. We estimate a life-cycle model incorporating fertility, labor supply and transfer decisions in which household decisions are modeled as a non-cooperative game focusing on the effects of parental time investment and parents' characteristics on children's life-time earnings market and marriage market outcomes. We also analyze the effect of the noncooperative decisions and allocation of resources within households on fertility and children's outcomes.

2.0 ESTIMATION OF DYNASTIC LIFECYCLE MODELS

The dynastic lifecycle model composes of two parts. The dynastic aspect of the model is the linkage between the individuals of different generations who faces the same problem in their own generations. The valuations of the future generations as the name dynastic suggests can affect the current generation individuals' choices and preferences. The choices are allowed to be sequential in a particular generation and this constitutes the lifecycle aspect of the problem. In general this model, not necessarily be limited to analyze the individual's problem across generations. Any framework fitting into this description can be estimated using the estimator developed in this chapter. However, given that the following chapter is using the estimator developed in this chapter, the estimation framework will be illustrated using an individual/household problem where the decisions of the current generation individual/household can affect their offsprings's outcomes, in fact can affect whether there will be offspring or not[1].

The chapter is organized as follows. Section 1 describes the dynastic lifecycle model to be estimated. The framework for the empirical implementation of the estimation is described in Section 2. Section 3 shows the estimation of the model. Section 4 compares the new estimator to the full solution estimator in a Monte Carlo study. Section 5 extends the results to a model with T periods in the lifecycle. Section 6 concludes. The appendix present proofs and implementation details.

[1]For instance fertility decision is part of the choices individual/household chooses.

2.1 MODEL

The model is an overlapping generation model with endogenous consumption, labor supply, fertility and investment in children decisions. The framework incorporates altruistic preferences, as in the Barro and Becker (1989)(9) model and builds on the literature which generalizes it (see Alvarez (1999)(5) , Enchevarria Merlo (1999)(30), Doepke (2004)(28) among others). We begin by describing the basic intergenerational problem and then extend it to a model which includes intra-generation life-cycle sequential decisions.

2.1.1 Individual Choices and Children's Outcomes

An adult from generation $g \in \{0, ...\infty\}$,live for T periods in which she makes decisions, $t \in \{0, 1, .., T\}$ which includes the birth decision $b_t \in \{0, 1\}$, and possibly transfers to children, $d_t \in \pi_h$. The transfer can be human capital or monetary transfer which affects the child's outcomes, but for simplicity we ignore bequest and transfers of assets and focus on transfers which affect the earnings potential of the child. We denote by N_t the total number of children at the beginning of period t. $D_t = \{d_{0,..}, d_{t-1}\}$ is a vector of transfers to children up to period t. Denote specific choice made in a particular period t by k_t where $k_t \in K$. Denote by $F(x_{t+1}|x_t, k_t)$ the stochastic transition function of the state variables, conditional on last period state variable and choice. We assume that all transition functions are known to all individuals in all periods and generations.

An individual's time invariant characteristics are denoted by x, it includes variables such as education, race and skill. We denote the children's time invariant characteristics by x'. The vector x_t denotes the persistent state variables at the beginning of period t; it includes x, N_t, D_t and belongs to the space $x_t \in s(x_t)$

The children's outcomes (their state variables) is a stochastic function of the parent's characteristics and her transfers. The production function of the child's characteristics is a stochastic function which depends on the individual's total transfers over the life cycle, D_s, where s indexes the child's year of birth. Denote the stochastic outcome function of a child born in period s by $M(x'|x, D_s)$.

We assume for simplicity that the transfer to each child is the same in the model, but it can be extended to include a gender-specific transfer (see Enchevarria and Merlo (1999)(30)). In our formulation, therefore, the parental time and monetary investments in children's education and productive skill affect the probability of their educational attainment and skills developed. In addition, it allows for the value of time investment in children to depend on parents' characteristics, such as education and skills through x.

We do not model explicitly the marriage decisions and do not make distinction between different genders of the individual for the sake of making the model simple for illustration of the estimator, such extensions are naturally included when the model is estimated using data.

2.1.2 Preferences

Each period there are preference shocks to the utility associated with each choice, denoted by $\varepsilon_t = [\varepsilon_{1t}, .., \varepsilon_{tK}]$; the shocks ε_{kt} are drawn independently across choices, periods and generations from a distribution function F_ε. The shocks are also conditionally independent (of all state variables). The individual per period utility depends on the current state x_t, whether there is a birth in that period and the preference shock ε_{kt}. The discount factor of the valuation of the children's utility is given by $\lambda N^{1-\nu}$, where N is the total number of children a person has at the end of the life cycle (at the end of period T). β is the annual discount factor. Denote by U_g the discounted expected lifetime utility of an individual in generation g at period 0

$$U_g = E_0 \left\{ \sum_{t=0}^{T} \beta^t \left[u(x_t, b_t, d_t) + \varepsilon_{kt} \right] + \beta^T \lambda \frac{N^{1-\nu}}{N} \sum_{t=0}^{T} b_t U_{g+1} \right\}. \tag{2.1}$$

The first element on the right hand side is the per period utility of an adult in generation g. The per-period utility also depends on whether there is a birth in the household capturing costs of birth, the number of children (which captures the reduction in consumption due to the costs of raising children and possibly a utility value of having the children). The second element is the altruistic component of the preferences; it captures the average expected lifetime utility of a child weighted by the discount, $\lambda N_\sigma^{1-\nu}$, which is assumed to be concave in the number of children, thus $0 < \nu < 1$.

Assumption: The problem is stationary across generations.

Stationarity means that the state $(x_t, x'_t \in s(x_t))$ and action spaces $(k_t, k'_t \in K)$ are same across generations and the utility and transition functions have the same functional form across generations. Therefore given that $x_t = x'_t$, $U_{gt} = U_{g+1,t}$ where U_{gt} is the period t counterpart of equation 2.1.

Under the assumption of stationarity, we can omit the generation index g. We first define the ex-ante value function V as the discounted sum of future utilities. This is the the discounted sum of future utilities for the individual before individual-specific preference shocks are observed and actions taken. Let's also define by $p(k_t|x_t)$ the conditional ex ante (again before ε_t is observed) probability that action profile k_t will be chosen conditional on state x_t. For $t < T$ the ex ante value function can therefore be written as

$$V(x_t) = \sum_{k_t} p(k_t = s|x_t) \left[u(k_t, x_t) + \beta \sum_{x_{t+1}} V(x_{t+1}) F(x_{t+1}|x_t, k_t) \right] \tag{2.2}$$

$$+ \sum_{s=1}^{K_t} p(k_t = s|x_t) E_\varepsilon[\varepsilon_t | k_t = s] \tag{2.3}$$

where E_ε denotes the expectation operator with respect to the individual-specific preference shocks.

Let $\upsilon(k_{jt}; x_t)$ denote individual's continuation value net of the preference shocks (also known as conditional valuation function) by choosing action k_{jt} conditional on the state variable x_t. This can be written as:

$$\upsilon(k_{jt}; x_t) = u(k_{jt}, x_t) + \beta \sum_{x_{t+1}} V(x_{t+1}) F(x_{t+1}|x_t, k_{jt}). \tag{2.4}$$

The choice k_{jt} is optimal if $\upsilon(k_{jt}; x_t) + \varepsilon_{jt} \geq \upsilon(k_{j't}; x_t) + \varepsilon_{j't}$ for all $k_{j't} \neq k_{jt}$. Thus, we can characterize the probability distribution over k_{jt} for all j and write the conditional ex ante choice probabilities of the choice profile:

$$p_{jt}(k_{jt}|x_t) = \int \left[\prod_{k_{\sigma jt} \neq k_{j'it}} 1\{\upsilon(k_{jt}; x_t) - \upsilon(k_{j't}; x_t) \geq \varepsilon_{jt} - \varepsilon_{j't}\} \right] dF_\varepsilon \tag{2.5}$$

where $\upsilon(k_{jt}; x_t) - \upsilon(k_{j't}; x_t)$ is the differences in the ex-ante conditional valuation when individual chooses k_{jt} and the valuations when $k_{j't}$ is chosen. Notice that the choices k_{jt} and

$k_{j't}$ are chosen such that k maps for every period state variables (x_t, ε_t) into choices, and we describe the probability distribution over the choices of an individual when the strategy is optimal.

Remember the intergenerational transition function of the persistent state variables of a child born in period s in the parent's life cycle is denoted by $M(x'|x_s, D_s)$. This function captures the stochastic outcomes of the child in terms of the child time invariant characteristics and the child's spouse characteristics, given the parents' time invariant characteristics and transfers to the child. The conditional valuation function in the final period of the life cycle T is given by

$$v(k_{jT}; x_T) = u(k_{jT}, x_T) + \beta\lambda \frac{(N_T + b_T)^{1-v}}{(N_T + b_T)} \overline{V}_N(k_{jT}; x_T) \tag{2.6}$$

Where $\overline{V}_N(x_T)$ is sum of the expected valuation over all children born up to period T plus the valuation of a child born in period T if there is birth

$$\overline{V}_N(k_{jT}; x_T) \equiv \sum_{s=0}^{T-1} \left[b_s \sum_{x'_0} V_s(x'_0) M(x'_0|x_s, D_s) \right] + b_T \sum_{x'_0} V_T(x'_0) M(x'_0|x_T, D_T) \tag{2.7}$$

Note that D_T and D_s for $s < T$ are both *functions of k_{jT}.* In the final period of the life cycle, the valuation function Equation 2.6 depends on current utility, and the discounted expected value of the children's valuation functions. The first element of Equation 2.7 is the expected valuation of the existing children at the beginning of period T, which state variables depend on past parental time input and the current period inputs. The second element is the expected value of a child born in period T for which the gender is unknown at the beginning of the period. Thus, this element depends on the birth decision and parental transfer. We assume that all children become adults after period T and their state variables are unknown until then regardless of the time of birth.

2.1.3 Representation

We use a representation of the valuation function in terms of the model's primitives and choice probabilities which allows for the estimation of the problem in several steps. Observe that the conditional valuation function for period t given in equation 2.4, the conditional valuation function for period T given in 2.6 and the definition of the dynastic component given in 2.7 can be used recursively to derive a representation for the period t conditional valuation function as follows (proof in appendix):

$$\begin{aligned} \upsilon(k_{jt}; x_t) &= u(k_{jt}, x_t) \\ &+ \sum_{s=t+1}^{T} \beta^{s-t} \sum_{x_s} \left\{ \left(\sum_{k_s} [u(k_s, x_s) + E_\varepsilon(\varepsilon_s | k_s = s)] p(k_s = s | x_s) \right) F(x_s | x_t, k_{jt}) \right\} \\ &+ \lambda \beta^{T-t} \sum_{x_0} V(x_0') H(x_0' | x_t, k_{jt}) \end{aligned} \tag{2.8}$$

where $F(x_s|x_t, k_{jt})$ is the $s-t$ transitions, $H(x_0'|x_t, k_{jt})$ is weighted generation transitions, and $V(x_0)$ is a vector of the ex-ante valuation functions . The transition function $H(x_0'|x_t, k_{jt})$ can be written as recursive function of $F(x_{t+1}|x_t, k_{jt})$, $M(x'|x, D_s)$, N_T, b_s, p and $1 - \nu$.

Define the ex-ante conditional lifetime utility as period t, exclusion the dynastic component as:

$$\begin{aligned} U(k_{jt}, x_t) &= u(k_{jt}, x_t) \\ &+ \sum_{s=t+1}^{T} \beta^{s-t} \sum_{x_s} \left\{ \left(\sum_{k_s} [u(k_s, x_s) + E_\varepsilon(\varepsilon_s | k_s = s)] p(k_s = s | x_s) \right) F(x_s | x_t, k_{jt}) \right\} \end{aligned}$$

Therefore we can write an alternative representation for the ex-ante value function as time t :

$$\begin{aligned} V(x_t) &= \sum_{k_{jt}} \left[U(k_{jt}, x_t) + E_\varepsilon(\varepsilon_{jt} | k_{jt}, x_t) \right] p_t(k_{jt} | x_t) \\ &+ \sum_{k_{jt}} \left[\lambda \beta^{T-t} \sum_{x_0} V(x_0) H(x_0 | x_t, k_{jt}) \right] p_t(k_{jt} | x_t) \end{aligned} \tag{2.9}$$

Equation (2.9) is satisfied at every state vector x_t, and since the problem is stationarity over generation at period 0 we express it as a matrix equation (proof in appendix):

$$\begin{aligned} V(X_0) &= P(X_0)U(X_0) + e(X_0, P(X_0)) + \lambda\beta^T P(X_0)H(X_0)V(X_0) \\ &= [I_{S(X)} - \lambda\beta^T P(X_0)H(X_0)]^{-1}[P(X_0)U(X_0) + e(X_0, P(X_0))] \end{aligned} \tag{2.10}$$

The terms on the right hand side of Equation 2.10 are the intergeneration and the per period discount factors, the choice probability matrix, the intergeneration state transition matrix, the ex-ante conditional lifetime utility, and the expected shocks. In matrix notation $V(X_0) = [V(x_0)]_{x_0 \in X_0}$ is $S(X_0) \times 1$ vector of expected discounted sum of future utility; $P(X_0)$ is $S(X_0) \times (S(K) \cdot S(X_0))$ dimensional matrix consisting of the choice probability $p(k|x_0)$ in rows x_0 and $S(X)$ and columns (k, x_0) and $(k, S(X))$, zeros in rows x_0 and $S(X)$ and columns (k, x'_0) and $(k, S(X))$ with $x'_0 \neq x_0$; $e(X_0, P(X_0))$is the $S(X_0) \times 1$ vector of expected preference shocks with element $[E_\varepsilon(\varepsilon_j|k_j, x)p(k_j|x)]'_{x \in X_0}$; and $I_{S(X)}$ denotes the $S(X_0)$-dimensional identity matrix. The second line in Equation (2.10) is a direct implication of the dominant diagonal property, which implies that the matrix $[I_{S(X)} - \lambda\beta^T P(X_0)H(X_0)]$ is invertible.

The representation of the dynastic component in equation (2.10) format allows replacing the term $V(x_0)$ in equation (2.9) by the derived representation in (2.10). Therefore this representation can be used to apply a Hotz-Miller type estimation algorithm to the intergenerational model introduced.

2.2 ESTIMATION

The difficulty of estimating the model is due to the non-standard nature of the problem. While the problem can be solved with a nested fixed point algorithm[2], it becomes computational intensive quickly, limiting the scope of the problem that can be analyzed. The alternative representation developed above of the continuation value of the intergenerational

[2]Equation (2.9) constitutes a fixed point problem in $V(x)$ and can be solved to obtain those functions for every possible value of x. For the solution of dyanamic discrete choice probalems with NFXP, see for instance Rust (1987) (73).

problem enables us to derive the necessary representation and apply the Hotz and Miller (1993)(43) estimation technique for single agent problems to the dynastic problem or a Pseudo-Maximum Likelihood estimator (i.e Aguirregabiria and Mira (2002)(1)).

Under the assumption that ε_s is distributed i.i.d. type I extreme value then conditional choice probabilities are related to the conditional valuation functions as follows:

$$\log\left(\frac{p_t(k_{jt}|x_t)}{p_t(k_{0t}|x_t)}\right) = \upsilon(k_{jt};x_t) - \upsilon(k_{0t};x_t) \tag{2.11}$$

for $k_{jt} \neq k_{0t}$. The i.i.d. type I extreme value assumption also implies that the conditional expectation of the preference shocks are functions of the choice probabilities as $E_\varepsilon(\varepsilon_t|k_{jt}) = \zeta - \log\left(p_t(k_{jt}|x_t)\right)$ where ζ is the Euler Constant (˜0.57721).

Estimation of the intergenerational model means using data on observable state variables and the actual choices made by agents, to obtain an estimate of the parameters of the functions: $u(x_t, k_t, \theta_u^0)$, $F(x_{t+1}|x_t, k_t, \theta_2^0)$, $M(x_0'|x_T, k_t, \theta_3^0)$, and the discount factors β^0, λ^0, and ν^0. Let $\theta^0 = (\theta_u^0, \beta^0, \lambda^0, \nu^0)$ denote the structural parameters of interest.

Suppose we have a data set which consists of a panel of observations from a random sample of decision makers in a particular generation g, $\{x_{it}, h_{it}, h_{iNt}, b_{it} : \; i = 1, ..., I, \;\; t = 0, T\}$, and a cross-section of observations for their successors in generation $g+1$ at $t = 0$, $\{x_{i0}' : \; i = 1, ..., I\}$. The representations developed in the previous section for the conditional value functions enables us to estimate the primitives of the model.

First we note that the transition functions $(F(x_{t+1}|x_t, k_t, \theta_2^0)$, $M(x_0'|x_T, k_t, \theta_3^0))$, and the conditional choice probabilities $p_t(k_{jt}|x_t)$ can be estimated directly from the data. Next, we use the relation in equation (2.11), to estimate the intergenerational model either by pseudo-maximum likelihood or GMM.

Suppose $\hat{\theta}_2$, $\hat{\theta}_3$ are consistent estimates of the parameters of the transition functions and $\left\{(\hat{p}(k_{jt}|x_t^l))_{l=1}^{S(x)}\right\}_{j=1}^{K}$ consistent estimates of the conditional choice probabilities $p_t(k_{jt}|x_t)$[3].

[3]For instance, they can be estimated consistently as cell estimators from data on choices and state variables.

2.2.1 Pseudo-Maximum Likelihood

Given that the data $\{x_{it}, a_{it}: \ i = 1, ..., I; \ t = 0, T\}$ were generated from the structural model with the parameters $(\theta^0, \theta_2^0, \theta_3^0)$, the pseudo-likelihood estimator is defined as :

$$\hat{\theta}_{PML} = \arg\max_{\theta} \left(\sum_{i=1}^{I} \sum_{t=0}^{T} \sum_{k_t=1}^{K} \mathcal{I}\{k_{it} = k_t\} \ln[p(k_t|x_{it};\theta)] \right)$$

where

$$p(k_{jt}|x_{it};\theta) = \int \mathcal{I}\{\upsilon(k_{jt}; x_{it}, \theta) + \varepsilon_{jt} > \upsilon(k_{j't}; x_{it}, \theta) + \varepsilon_{j't} \ \forall k'_{jt} \neq k_{jt}\} dF_{\varepsilon} \tag{2.12}$$

and $\upsilon(k_{jt}; x_{it}, \theta)$ are constructed using the choice probabilities and the transition functions which are estimated in an earlier step[4].

2.2.2 GMM

Under the assumption that ε_s is distributed i.i.d. type I extreme value then Hotz and Miller inversion implies that

$$\log\left(\frac{p_t(k_{jt}|x_t)}{p_t(k_{0t}|x_t)}\right) = U(k_{jt}, x_t) - U(k_{0t}, x_t) + \lambda\beta^T \sum_{x_0} V(x_0)[H(x_0|x_t, k_{jt}) - H(x_0|x_t, k_{0t})] \tag{2.13}$$

for $k_{jit} \neq k_{0it}$.Define the $(K-1) \times 1$ vector $\xi_{it}(\theta)$ (which is obtained by subtracting the right hand side of equation (2.13) from the left hand side for $k_{jt} = 2, ..K$) as the vector of moment restrictions for individual i for period t as $\xi_{it}(\theta) = (\xi_{it2}(\theta), ..., \xi_{itK}(\theta))'$. Define the $[(T+1) \times (K-1)] \times 1$ vector $\xi_i(\theta) = (\xi'_{i0}(\theta), ..., \xi'_{iT}(\theta))'$ as the vector of moment restrictions for a given individual over time. The $[(T+1)\times(K-1)]\times[(T+1)\times(K-1)]$ weighting matrix $\Xi_i(\theta)$ is defined as $\Xi_i(\theta) \equiv E\left[\xi_i(\theta)\xi'_i(\theta)\right]$.Notice that the matrix $\Xi_i(\theta)$ is block diagonal with diagonal elements defined as $\Xi_{it}(\theta) \equiv E_t[\xi_{it}(\theta)\xi'_{it}(\theta)]$, and off-diagonal elements that are zero because $E_t[\xi_{it}(\theta)\xi'_{is}(\theta)] = 0$ for $s \neq t$. The $(K-1) \times (K-1)$ conditional heteroskedasticity

[4] In this estimation, the conditional valuation functions from the estimated choice probabilities and the transition functions are used to obtain the choice probabilities again from the realtion given in (2.11). In this sense it is like a one step iteration on the choice probabilities starting from the estimated consistent ones from the data.

matrix $\Xi_{it}(\theta)$ associated with the individual-specific moment restrictions ξ_{it} is evaluated using an initial consistent estimator of θ^0. The optimal GMM estimator for θ satisfies:

$$\hat{\theta}_{GMM} = \arg\min_{\theta} \left(\frac{1}{I} \sum_{i=1}^{I} \xi_i'(\theta) \widehat{\Xi}_i^{-1}(\theta) \xi_i(\theta) \right)$$

The individual moment restrictions at period t, $\xi_{it}(\theta)$ are formed by introducing error in evaluating the sample counterparts of the moment conditions. The particular element ξ_{itk_t}, which is the restriction for choice k_t is calculated as follows:

$$\xi_{itk_t}(\theta) = q_{k_t}(\hat{p}(1|x_{it}), ...\hat{p}(K|x_{it})) - (\upsilon(k_t; x_{it}, \theta, \hat{\theta}_2, \hat{\theta}_3) - \upsilon(1; x_{it}, \theta, \hat{\theta}_2, \hat{\theta}_3)) \quad for \quad k_t = 2, ..K_p$$

where $q_{k_i}(\hat{p}(1|x_{it}), ...\hat{p}(K_p|x_{it}))$ is the inverse distribution function as defined in Hotz and Miller (1993) and it is . The estimated parameter vector $\hat{\theta}_{GMM}$ is a consistent estimator of the true parameters θ^0.

2.3 NUMERICAL EXAMPLE AND MONTE CARLO STUDY

In order to compare the dynamics of the model in a numerical example and examine the performance of the estimator, we use a simple human capital investment model with intergenerational transfers which has the two period model structure of Section 1. First we generate simulated data from the model for given parameter values, compare the dynamics and then estimate the model parameters for the generated dataset. We obtain ML estimates using the NFXP (Nested Fixed Point) and PML (Pseudo Maximum Likelihood) estimates using our estimator. The estimations are repeated for both algorithms for different specifications of the model in terms of sample size (i.e., for 1000, $10,000$, $20,000$, $40,000$). The number of structural parameters estimated including the discount factors are 3.

Table 1: State Transition Matrix

c	$k_0 = 0$					$k_0 = 1$				
	0.5	**0.6**	**0.7**	**0.8**	**0.9**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0.5	0.85	0.13	0.02	0	0	1	0	0	0	0
0.6	0.04	0.85	0.09	0.02	0	0.1	0.9	0	0	0
0.7	0.01	0.04	0.85	0.09	0.01	0.13	0.27	0.6	0	0
0.8	0	0.01	0.05	0.85	0.09	0.01	0.11	0.28	0.6	0
0.9	0	0	0	0	1	0	0.04	0.13	0.23	0.6

2.3.1 Model Environment

The period utility function has the following linear form where agent chooses whether to invest or not $k_t \in \{0, 1\}$ in each period $t \in \{0, 1\}$. She gets the following utilities associated with each choice:

$$u(c_t, k_t, \varepsilon_t) = \left\{ \begin{array}{cc} c_t + \epsilon_t(0) & \text{if } k_t = 0 \\ (1-\theta)c_t + \epsilon_t(1) & \text{if } k_t = 1 \end{array} \right\}$$

where $\varepsilon_t(k_t)$ is the choice specific unobservable part of the utility and assumed to be i.i.d. extreme value type I.

In the example environment it is assumed that each agent starts the lifecycle with a particular consumption value $c_t \in (0.5, 0.6, 0.7, 0.8, 0.9)$. The transition from one state to another is probabilistic and denoted by the transition matrix $F(c_1 \mid c_0, k_0)$, which is given in Table 1.

The next generation's starting consumption value c' depends on the sum of the investment decisions in the life-cycle, where $D \in (0, 1, 2)$. This transition is governed by the intergenerational transition function $M(c'_0 \mid D)$ given in Table 2, where c'_0 is the consumption of the next generation at period 0.

The transition is such that if the agent opts to invest 2 times in the life-cycle, then she can increase the probability that the next generation will start his lifecycle with the

Table 2: Intergenerational Transition Matrix

c'_0 :	0.5	0.6	0.7	0.8	0.9	D
	1	0	0	0	0	0
	0	0.1	0.4	0.4	0.1	1
	0	0	0.04	0.06	0.9	2

highest consumption level to 0.9. Next generation's starting consumption level will be the determined by the probabilities given in the row corresponds to the investment level. If she invests nothing, then the next generation will have the lowest consumption value. Each row corresponds to one of the values of $D \in (0, 1, 2)$ where the first row is for the investment $D = 0$.

2.3.2 Results

First we simulated the model for a given values of the parameters of the model. $(\theta, \lambda, \beta) = (0.25, 0.8, 0.95)$. where θ is the structural parameter of interest which gives the marginal cost of investment. λ and β are the generation and time discount factors respectively. We produced samples of $1,000$, $10,000$, $20,000$, $40,000$ observations for 100 samples. For the PML estimation, the initial consistent estimates of the CCPs are estimated nonparametrically using the generated sample. Next we estimated the model by NFXP and PML[5]

Table 3 presents the result of the estimations for each specification. The mean, standard deviation, bias and Mean Squared Error (MSE) of each parameter estimate are reported in the respective column for each sample size. The bias and the MSE are calculated relative to the original DGP (Data Generating Value) value of the parameter. The DGP value of the parameter is also reported at the top left corner of summary statistics block for that parameter. We find that the finite sample properties of the estimators improve monotonically

[5] As illustrated in the estimation section, intergenerational models at the final step can be estimated either by PML or GMM. For this simulation study we used the PML.

Table 3: Simulation Results

The estimated parameter values and their computation time. Pseudo Maximum Likelihood (PML) corresponds to the estimation conducted by the new estimator using PML and ML estimation is by the Nested Fixed Point (NFXP).

	Pseudo Maximum Likelihood sample size(n)				Nested Fixed Point (ML) sample size(n)			
$\theta = 0.25$	1,000	10,000	20,000	40,000	1,000	10,000	20,000	40,000
mean	0.24473	0.24935	0.24886	0.24881	0.22714	0.24571	0.23320	0.24477
stdev	0.04991	0.01328	0.00915	0.00668	0.04884	0.01354	0.02135	0.01019
bias	-0.00527	-0.00065	-0.00114	-0.00119	-0.02286	-0.00429	-0.01680	-0.00523
MSE	0.00249	0.00017	0.00008	0.00005	0.00288	0.00020	0.00073	0.00013
$\lambda = 0.8$								
mean	0.80425	0.79745	0.79797	0.79673	0.77538	0.78966	0.76934	0.78855
stdev	0.11241	0.03175	0.02157	0.01587	0.09211	0.03244	0.03656	0.02063
bias	0.00425	-0.00255	-0.00203	-0.00327	-0.02462	-0.01034	-0.03066	-0.01145
MSE	0.01253	0.00100	0.00046	0.00026	0.00901	0.00115	0.00226	0.00055
$\beta = 0.95$								
mean	0.94208	0.95245	0.95037	0.95136	0.93441	0.95227	0.94603	0.95027
stdev	0.06276	0.01893	0.01301	0.00934	0.05322	0.01983	0.01820	0.01236
bias	-0.00792	0.00245	0.00037	0.00136	-0.01559	0.00227	-0.00397	0.00027
MSE	0.00396	0.00036	0.00017	0.00009	0.00305	0.00039	0.00034	0.00015
Avg Comp time[6]	0.65	2.88	6.06	12.60	347.6	376.4	467.5	509.8

with sample size. For the NFXP, MSE drops more than 10 times when moving from sample size of $1,000$ to $10,000$, and drops more than 6 times when moving from $n = 20,000$ to $n = 40,000$ for the parameter θ. The results for the discount factors are similar. The MSE drops approximately 8 times when moving from sample size of $1,000$ to $10,000$, and drops 4 times when moving from $n = 20,000$ to $n = 40,000$ for the parameter λ. The reduction in MSE for moving from $1,000$ to $10,000$ is 7 times, and it is 2 times when we move from $n = 20,000$ to $n = 40,000$. We observe similar patterns for PML.

For the sample size of $1,000$, we obtain MSE: 0.00249 compared to 0.00288 for θ, 0.01253 compared to 0.00901 for λ and 0.00396 compared to 0.00305 for β from the PML estimator compared to NFXP. For the sample sizes $10,000, 20,000$ and $40,000$, MSE obtained from PML is lower, however the magnitudes are quite close. In terms of biases, the two estimation algorithms are quite similar. However, the two estimation algorithm differs greatly in terms of computation times. The average computation time for the NFXP for $n = 1,000$ is 347.6 seconds compared to only 0.65 seconds for *the PML.* The ratio is 530. For the sample size of $40,000$ computation times are 509.8 and 12.6 respectively with a ratio of 40.4[7].

2.4 CONCLUSION

This paper develops a framework for estimation of life-cycle dynastic models with altruistic preferences. We develop an alternative representation of the continuation value of the intergenerational problem which enables us to estimate the model in multiple steps using a CCP estimator. The estimator can be used to estimate a large class of dynastic models with endogenous choices within the dynasty. Moreover the framework allows the agent in a dynasty to sequence his choices in a finite number of periods. This estimation framework encompasses the lifecycle dynastic models which we use in Chapter 3, however it is applicable to any framework fitting the description. We illustrated the estimator using a version of a lifecycle dynastic model which is essential for the application in the following chapter.

[7] Calculation of the computation times does not include the cases where the NFXP algorithm fails to converge. Especially for the sample size of $1,000$, we had to either change the convergence criteria or the seed used in constructing the random sample in approximately 15% of the 100 replications. We encountered similar convergence problems in sample sizes $10,000$, 20.000 and $40,000$, but less often.

The finite sample performance of the new estimator is compared to the full solution estimator (NFXP), and proves to have comparable finite sample properties while reducing the computational time considerably. At least within the field of labor economics as a start, the estimator will allow the intergenerational models including generational transfers estimable which previously could only be analyzed theoretically or numerically by calibration exercises.

3.0 ESTIMATING THE PARENTAL RETURNS TO TIME INVESTMENT IN CHILDREN

Full Title: **Estimating the returns to parental time investment in children using a lifecycle dynastic model**

Co-authored with George-Levi Gayle and Limor Golan

Parental investment in children plays an important role in the intergenerational persistence of earnings. This paper estimates the returns to parental time input in children of parents with different characteristics and across demographic groups. In order to quantify these returns, we develop a model of dynastic households in which altruistic individuals choose fertility, labor supply, and time investment in children sequentially. Using data on two generations from the PSID, this framework enables us to estimate the costs and returns of time investment in children.

There is an extensive empirical literature showing that parental inputs and characteristics are important determinants of children's achievements measured by short-term outcomes such as test scores, (see Todd and Wolpin (2003)(81), Cunha and Heckman(2008)(23) among others) and long-term outcomes such as completed education and labor market outcomes. For example, Berman, Foster, Rosenweig and Vashishtha (1999)(15) provide evidence on the effect of schooling of mothers in India on their children's schooling outcomes (see also Rosenweig and Wolpin (1994)(71) for a study using NLSY data, and Black and Devereaux (2011)(16) for a survey on the literature). Studies in this literature can be divided into those that use family background variables as a proxy for parental input and those that provide direct evidence on the effect of parental time investment in children on their educational and cognitive outcomes. See Murnane, Maynard, and Ohis (1981)(64), Guryan, Hurst and Kearney (2008)(37), Datcher-Loury (1988)(24), Houtenville and Smith Conway (2008)(45), Leibowitz 1974, 1977(51)(52), Hill and Stafford 1980(40), Kooreman and Kapteyn 1987(49) for examples of studies using the direct approach (see Juster and Tafford (1991)(47) for a survey on empirical evidence of time allocation).

We contribute to the literature by measuring the returns to parental time investment in a life-cycle dynastic model in which fertility, labor supply and time spent with children decisions are endogenous. In contrast to previous studies, the returns to investment are measured in terms of the children life-time utility. As documented in the literature, investment in children varies substantially with family demographic characteristics and wealth. By modeling labor supply and time investment choices, we are able to explicitly account for the impact of households characteristics on investment in children. Specifically, we account for heterogeneity (i.e. differences in education, parents skills, family structure and race) in the costs and in the returns on parental time investment. The costs are measured in terms of decrease in leisure and loss of labor market earnings. The returns are measured by the impact of parental time input on educational attainment of children, their skills and therefore life time earnings, as well as their marriage market outcomes; all these factors are aggregated and measured in terms of expected life-time utility of children. In addition, there is substantial variation in investment in children across household with different number of children. By modeling fertility choices, we capture the quantity-quality trade-off that households with different demographic characteristics face.

Models of dynastic households have been traditionally used to analyze investment in children and persistence in earnings and wealth across generations (e.g. Loury (1981)(53) Laitner (1992)(50) and the work by Becker and Tomes (1979), (1986)(12)(13) on parental time investment in children). A second class of dynastic models, pioneered by Becker and Barro (1988)(8) and Barro and Becker (1989)(9) analyzes fertility decisions and transfers to children. A small number of empirical paper quantify the returns to parental investment in children using dynastic models. Rios-Rull and Sanchez-Marcos (2002)(70) studies the returns of parental investment in children's education, their earnings and marriage market, Doepke and Tertilt (2009)(29) allows the returns on investment in children's human capital to depend on the parents' education and Echevarria and Merlo (1999)(30) in which a dynastic model of household bargaining gives rise to a gender gap in parental investment in education of the children. Our paper contributes to this literature by using data on time investment in children and by incorporating life cycle into the Becker-Barro(9) framework, thus capturing the dynamic aspects of labor supply decisions, time investment in children and fertility.

To the best of our knowledge only two other papers estimate the returns to parental time investment in children in a life-cycle framework accounting for endogenous labor supply and the opportunity cost of parental time. Kang (2010)(48) estimates a life-cycle model with endogenous parental transfers, fertility and labor supply. In her paper parents derive utility from the quality of children measured by their education and skill which proxy for children wages. Similar to our model, parental time investment affects the educational outcome and a labor market skill of children. The main difference from our paper is that we use a dynastic model, thus measuring the returns in terms of children life time utility which aggregates explicitly the labor market returns, the marriage market returns, and the utility derived from their choices. In addition, we use data on parental time input while Kang (2010)(48) uses labor supply data as a proxy for parental time investment and focused on the impact of dissolution of marriage on the outcome of children. Del Boca, Flinn and Wiswall (2010)(27) also use data on time investment in children in a life-cycle model with endogenous labor supply and time investment in children. They measure the effect of time investment in children on unobserved quality of a child using data on test scores of children. Our contribution is different in several respects. As discussed above, we measure the effect of parental time investment on life-time utility of children. In addition, their paper estimates the returns using data on families with one child, thus we further contribute to this literature by modeling fertility choice and estimate the returns and quality-quantity trade-offs in households with multiple children.

In our framework individuals may be single or married, and divorce and marriage evolve according to a stochastic process, thus individuals may live in different households over the life cycle. In the literature, households decisions are either framed as a single decision maker problem (this approach is pioneered by Becker (1965)(7)) or as a bargaining problem which is either modeled as a cooperative game theoretic problem or as a non-cooperative one (e.g., Manser and Brown (1980)(56), McElroy and Horney (1981)(57), Chiappori (1988)(21); see also Chiappori and Donni (2009)(22) for a recent survey on non-unitary models of household behavior, and Lundberg and Pollak (1996)(54) survey on non-cooperative models of allocation within households). We model household decision problem as a noncooperative game and solve for a Markov Perfect Equilibrium (for models of household allocations which are determined as a Nash Equilibrium outcomes of a non cooperative game see Del Boca and Flinn (1995, 2010)(25)(26), and Chen and Wolley (2001)(20)). While there is no consensus in the literature regarding the process governing household decisions, there are several advantages to this approach in our framework. First, the Becker-Barro model is formalized as a single decision maker dynamic optimization problem. Since we solve for a Markov Perfect equilibrium, given any spouse strategies and characteristics, the problem reduces to a single agent optimization problem and fits naturally in their theoretical framework as well as in the estimation framework of dynamic games which we discuss below. At the same time, in contrast to a unitary model approach, we are able to evaluate separately the value function of each individual, which is an advantage as parents utility is derived from their own children utility and not from the utility of their spouse. Second, since individuals may belong to different households over their life cycle, and since parents care about the utility of their own children, formulating the optimization problem as an individual decision maker simplifies the representation and estimation of the problem relative to a household cooperative bargaining problem is more straightforward.[1]

In the model, each individual from each generation lives for T periods. Over the life-cycle, each individual makes labor supply and time investment decisions in children every period; only females make birth decisions every period. Marriage and divorce evolve according to a stochastic exogenous process. If there are two individuals in the households the decisions are modeled as a non cooperative game and are made simultaneously. We do not model explicitly bargaining over allocation of consumption within the households and assume that each individual receives (per period) utility from his own income, the spouse's income and the stock of existing children in the household. This formulation is consistent with transfers of income between spouses in which the size of the transfers depends on the number of

[1] To the best of our knowlegdge no paper has fully estimated a dynastic model with Nash bargaining solution, divorce and marriage. Echevarria and Merlo (1999) estimate implications of dynastic model with endogenous fertility in which household allocation is determined by a Nash bargaining solution in a model with no divorce and marriage.

children and earnings of each individual in the household. The total time investment in children of both spouses over the life cycle affects the children's outcomes through several channels. Once children become adults, their education levels are realized; the education level is a stochastic function of the parental time input and the parents education level and labor market skills. In addition, the skill level of a child and the education level of the child's spouse are a stochastic function of the child's education. Thus, parental time input and characteristics affect marriage outcomes and labor market skill indirectly. Therefore, although marriage is exogenous, parents take into account the marriage market outcomes of the children when they make investment and birth decisions.

The Becker-Barro framework provides a natural way to aggregate the value of the different aspects of the outcomes of the children by measuring the returns in terms of the discounted valuation function of the child. Time investment in children involves trading off leisure and hours worked in the labor market. Earnings are the marginal productivity of the individual and depend on the skill level, education, current level of labor supply and actual labor market experience. Thus, the opportunity costs of time includes current earnings as well as future loss of earnings resulting from accumulating less experience. This formulation allows us to capture the heterogeneity in the opportunity costs of time of parents by education, skill, race and gender groups. Because both the returns in terms of children outcomes and the opportunity costs of time depend on the parents productive characteristics the model can potentially generate decline in fertility for high earnings households (see Jones, Schoonbroodt and Tertilt (2008)(46) for discussions on fertility models).

We use a partial solution estimation method which is a modified version of the multistage estimation procedure developed in Gayle, Golan and Soytas (2010)(35). It uses the assumption of stationarity across generations and the discreteness of the state space of the dynamic programming problem to obtain an analytic representation the valuation function. This representation is a function of the conditional choice probabilities, the transition function of the state variable, and the structural parameters of the model. The conditional choice probabilities and the transition function are estimated in a first stage and used in the generation valuation representation to form the terminal value in the life-cycle problem. The life-cycle problem is then solved by backward induction to obtain the life-cycle valuation functions. Because the game between spouses is a complete information game, a sufficient condition for the existence of equilibrium in pure strategies is super modularity. Our game is super modular if there are strategic complementarities in time investment of parents or outcome of parental time investment is independent of the spouse's investment. An additional advantage of using a multiple step estimation approach is that it allows us to estimate the children's education production function parameters separately, using a Three Stage Least Square method, and verify that the conditions for existence of equilibrium are satisfied. We then form moment conditions from the best response functions and estimate it in a third step. Finally to reduce the computational burden of the backward induction in the life-cycle problem we use the forward simulation technique developed in Hotz, Miller, Sanders and Smith (1994)(44), and estimate the remaining structural parameters using Generalized Methods of Moment (GMM) estimator. To the best of our knowledge this is first paper to estimate a dynamic complete information game.

Our preliminary analysis shows that parental investment in children varies significantly across gender, race, education levels, and household composition. It also shows that after controlling for gender, education levels, and household composition, the differences across race are significantly reduced. We find that both maternal and paternal time investment increase the likelihood of higher educational outcome of their children. However, the impact is complementary; fathers' time investment increases the probability of graduating from high school and getting some college education while mothers' time increases the probability of achieving a college degree. The estimates of the education production-function show that girls have a higher likelihood than boys of achieving high levels of education, and that blacks have higher variance than whites in their educational outcomes, after controlling for parental inputs. Specifically, blacks have a higher probability of not completing high school than whites, however, they also have a higher probability of graduating from college than whites.

We then quantify the returns to parental time investment using the effect of an increase in time input on the change in the valuation function of the child. We find that the overall returns to fathers' time investment is only 40% that of mothers' time investment for white. We find the black mother's time investment is insignificant and the effect of time investment is only important for white mothers. Although both parents input improve the educational attainment of children, maternal time investment increases the probability of a child graduating from college, and a college degree increases the returns in both the labor and the marriage markets. Similar to Rios-Rull and Sanchez-Marcos (2002)(70), we find that both parents education levels, all else equal, increases the outcomes of the children but the effect of fathers' education is higher than the effect of mothers' education. There are race differences in the returns to paternal time investment and this interacts with the gender composition of the children in the household for both black and white fathers. One reason for insignificant maternal time investment by blacks may be the family structure. There is a significantly higher proportion of black single mothers than white single mothers and the opportunity costs of time for single mothers are higher than the opportunity costs of married mothers. Finally the returns to maternal time investment is independent of the gender of the child, whereas paternal time favors girls. This implies that fathers act in a achievement maximizing manner, favoring high ability children in the family. Since girls already have a higher likelihood of achieving high education outcome than boys, fathers seems to investment more time in girls than in boys as the number of children increases.

Our findings suggest a significant quality-quantity trade-off. This trade-off is measured in terms of the rate of increase in utility of parents versus the rate of the decline in the average life time utility per child resulting from having an additional child. The level of investment per child is smaller the larger the number of children, thus, this decline in the per child investment is driven by the time constraint and the opportunity costs of time and not by the properties of the production function technology of children. The negative relationship between income (education) and fertility is therefore explained by the higher opportunity cost of time of educated parents in terms of forgone earnings. We find similar quality-quantity trade-off for blacks and whites after controlling for education and parental inputs. Therefore the black-white gap seems to be related to the factors as education and the time investment of the parents when the returns are measured as the aggregate measure of utilities of future generations. This explanation is in line with Chiswick (1988)(19) evidence for quantity-quality trade-off; he concludes that family decisions and intergenerational transfers may play a big role in the observed race gap in achievements and earnings. Neal (2006)(65) provides evidence for the importance of these factors in the observed Black-White skill gap and its trends. Our direct estimates support this hypothesis.

Interestingly, we find that females have higher valuation functions (i.e. female child value is higher than that of a male child). Despite the fact that females earn less than men with the same productive characteristics, females are more likely to obtain higher levels of education than males, given equal amount of parental inputs and education is highly compensated in the labor market. However, even given the same level of education the valuation function of females are higher than males; this is because married females receive significant transfers from their husband's income. This findings can be explained by the fact than females are endowed with the birth decisions and males value children, but cannot make decisions to have them. This explanation is consistent with Echevarria and Merlo (1999)(30) which finds that transfers made within households increase the returns to parental investment in girls, and that the gender gap in education outcome of children is smaller when considering endogenous investment of parents in children.

The rest of the paper is organized as follows. Section 2 describes our data and variable construction. It also presents our preliminary analysis. Section 3 presents our theoretical model. Section 4 presents our estimation technique and empirical implementation. Section 5 presents the estimation results. Section 6 presents our measures of the quality-quantity trade-off and the return to parental time investment. Section 7 summaries our findings and concludes. Proofs are given in the appendix.

3.1 DATA

We used data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID). We selected individuals from 1968 to 1996 by setting the individual level variables "Relationship to Head" to head or wife or son or daughter. We dropped all sons or daughters if they are younger than 17 years of age. This initial selection produces a sample of 12,051 and 17,744 males and females respectively; these individuals were observed for at least one year during our sample period. Our main sample contains 423,631 individual-year observations.

We only kept white and black individuals between the ages of 17 and 55 in our sample. The earnings equation requires the knowledge of past 4 participation decisions in the labor market. This immediately eliminates individuals with less than 5 years of sequential observations. This reduces the number of individual-year observations to 139,827. In order to keep track of parental time investment throughout a child's early life we dropped parents we only observed after their children are older than 16 years of age. We also dropped parents with missing observations during the first 16 years of their children's life. Furthermore, if there are missing observations on the spouse of a mare individual then that individual is dropped from our sample.

The PSID measures annual hours of housework for each individual, however, it does not provide data on time parents spend on child care. This variable is estimated using a variation of the approach use in the previous literature. Example of papers using this approach can be found in Hill and Stafford (1980)(40), Leibowitz (1974)(51), and Datcher-Loury (1988)(24). Hours with children are computed as the deviation of housework hours in a particular year from the average housework hours of married individuals with no child by gender and education. Negative values are set to zero and child care hours are also set to zero for individuals with no children.

Table 4 presents the summary statistics for our sample; Column (1) summarizes the overall sample, Column (2) focuses on the parents, and Column (3) summarizes the characteristics of the their children. It shows that the first generation is on average 7 years older than the second generation in our sample. As a consequence a higher proportion are married in the first generation relative to the second generation. The male-female ratio is similar across generations (about 55 percent female), however, our sample contains a higher proportion of blacks in the second generation that in the first generation (about 29 percent in the second and 20 percent in the first generation). This higher proportion of blacks in the second generation is due to the higher fertility rate among blacks in our sample. There are no significant differences across generations in the years of completed education. As would be expected, because on average the second generation in our sample is younger that the first generation in our sample, the first generation has higher number of children, annual labor

Table 4: Summary Statistics

	(1)		(2)		(3)	
Variable	N	Mean	N	Mean	N	Mean
Female	115,280	0.545	86,302	0.552	28,978	0.522
Black	115,280	0.223	86,302	0.202	28,978	0.286
Married	115,280	0.381	86,302	0.465	28,978	0.131
Age	115,280	26.155	86,302	27.968	28,978	20.756
		(7.699)		(7.872)		(3.511)
Education	115,280	13.438	86,302	13.516	28,978	13.209
		(2.103)		(2.138)		(1.981)
Number of children	115,280	0.616	86,302	(0.766)	28,978	0.167
		(0.961)		(1.028)		(0.507)
Annual labor income	114,871	16,115	86,137	19,552	28,734	5,811
		(24,622)		(26,273)		(14,591)
Annual labor market hours	114,899	915	86,185	1078	28,714	424
		(1041)		(1051)		(841)
Annual housework hours	66,573	714	58,564	(724)	8,009	641
		(578)		585		(524)
Annual time spent on children	115,249	191	86,275	234	28,974	63.584
		(432)		(468)		(259)
Number of individuals	12,318		6,813		5,505	

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID), and include individuals surveyed between 1968 and 1997. Column (1) contains the summary statistics for the full sample; column (2) contains the summary statistics for the parents generation; column (3) contains the summary statistics of the off spring of the parents in column (2). Annual labor income is measured in 2005 dollars. Education measures year of completed education. There are less observations for annual housework hours than time spent on children because single individuals with no child are coded as missing for housework hours but by definition are set to zero for time spent on children.

income, labor market hours, housework hours, and time spent with children. Our second generation sample does span the same age range, 17 to 55, as our first sample.

3.1.1 Preliminary Analysis

Many studies have analyzed various dimensions of the relationship between mothers' time with children and children's outcomes (see Hill and Stafford (1980)(40), Leibowitz (1974)(51), Datcher-Loury (1988)(24), among others). Few studies, however, have analyzed the effect of fathers' time with children or household labor market decisions on their children's subsequent outcomes. In this section we document some of these empirical regularities as a way of motivating and clarifying our modelling choices.

3.1.1.1 The Relationship between Child Care Time and Household Composition

Figure 1 presents the kernel estimates of the density of hours spent with children by marital status, gender, and race. It shows that females provide significantly more hours than males, confirming the well documented specialization by gender in home production. The upper left hand panel shows that over the nonzero range, the distribution of hours spent with children does not differ significantly by marital status, however, there is a higher incidence of zero hours spent with children for married parents than for single parents. A closer look at the middle and bottom left hand panels shows that this higher incidence of zero hours with children for married parents versus single parents is mostly is due to the significantly higher incidence of zero hours among married versus single male parents. The middle left hand panel shows that the distribution, for time investment in children greater than 160 hours per annum, is similar across marital status for male parents. Below 160 hours per annum, married male parents are less likely to provide time with children than single male parents. Married female parents are more likely to provide high hours and are less likely to provide low hours than single female parents.

The right hand panels of Figure 1 present the distributions of child care hours by race and gender; they show that there are little to no differences in the distribution of hours spent with children of black and white parents. If anything, blacks provide more hours than whites. The pattern for the overall distribution by race is repeated for males, however, white females provide more hours than their black counterparts. This could be due to the higher incidence single mothers among blacks than whites; this is demonstrated by the similarity between the whites versus blacks' distributions and married versus single distributions for mothers. Figure 2 presents the kernel estimates of the density of hours invested in children by own education, spouse education, number children, and gender. The top panels show that fathers hours are increasing with fathers' education, with college educated fathers having the highest likelihood of providing time with children. However, the distributions of hours of mothers are not monotone in mothers' education; a mother with less than a high school education is most likely to provide high hours while a mother with some college education is least likely to provide high hours. The patterns observed for own education are repeated for spouse education, with the differences that a mother whose spouse has a college education is the least likely to provide high hours. This highlights the assortative mating on education in the marriage market. The bottom panels of Figure 2 present the distributions by the number children and show that hours provided by both fathers and mothers are increasing in the number of children.

3.1.1.2 The Relationship between Child Care Time and Labor Market Time

Time not spent taking care of children can either be spent working in the labor market or on leisure; given a fixed hours endowment day, it suffices to analyze the relationship between time investment in children and labor market time. Figure 3 presents the kernel estimate of the densities of hours spent with children by labor supply, education, and gender. The top panels of Figure 3 show that for both fathers there is a negative relationship between hours worked and hours spent with children. This may indicate some degree of substitutability between time with children hours provided by parents and market purchased child care. The second panels from the top of Figure 3 show that among parents who are not currently employed college graduates are more likely to spend more hours with children. Parents who did not complete high school and those that have some college education but not a college degree are the least likely to spent time with children on child when they are not working. Surprisingly, the behavior of parents with some college is similar to those with less than high school; this may reflect some selection on unobservable which are correlated with not completing a given level of education. We seek to capture these unobserved traits by using individual specific effects that are correlated with observed individual specific variable such as the level of completed education. The third panels from the top show that this pattern is repeated for parents that are currently working part-time. The bottom panels of Figure 3 show that these patterns are very different for parents that are working full-time in the labor market. For fathers that are working full-time in the labor market there are virtually no differences by education groups; however, for mother working full-time those with less than high school education are more likely to spend a high number of hours with children. On the other hand, mothers that have at least a college degree are the least likely to spend a large amount of hours with children when they are working full-time. This may reflect differences in the type of full-time jobs perform by mother with at least a college education and mothers with less education. Nevertheless, these empirical findings demonstrate the interplay between time investment in children, gender, education, household composition, and the labor market hours.

Figure 1: Parental Time Densities by Marital Status, Gender and Race

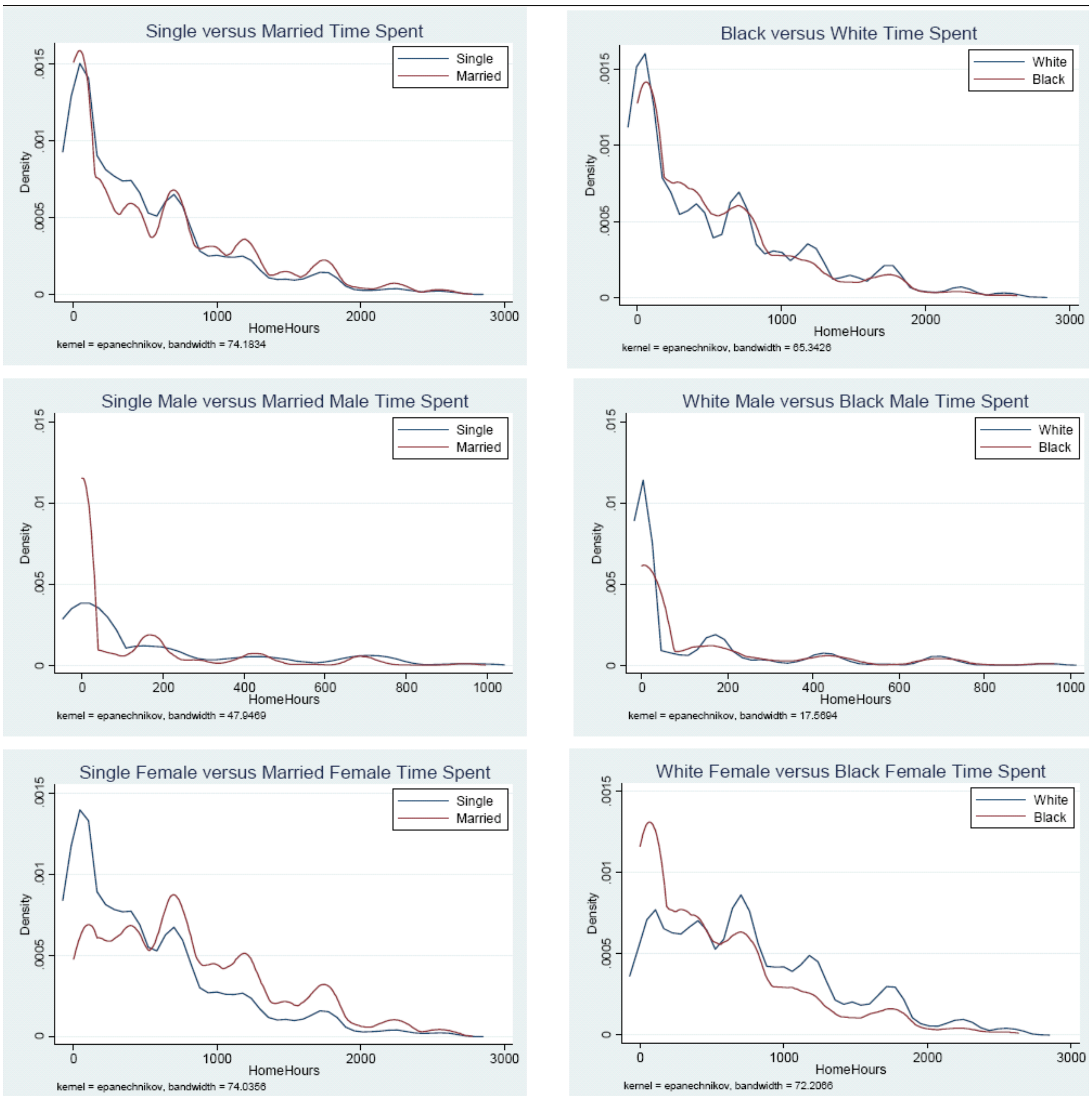


Figure 2: Parental Time Densities by Own Education, Spouse's Education and Number of Children

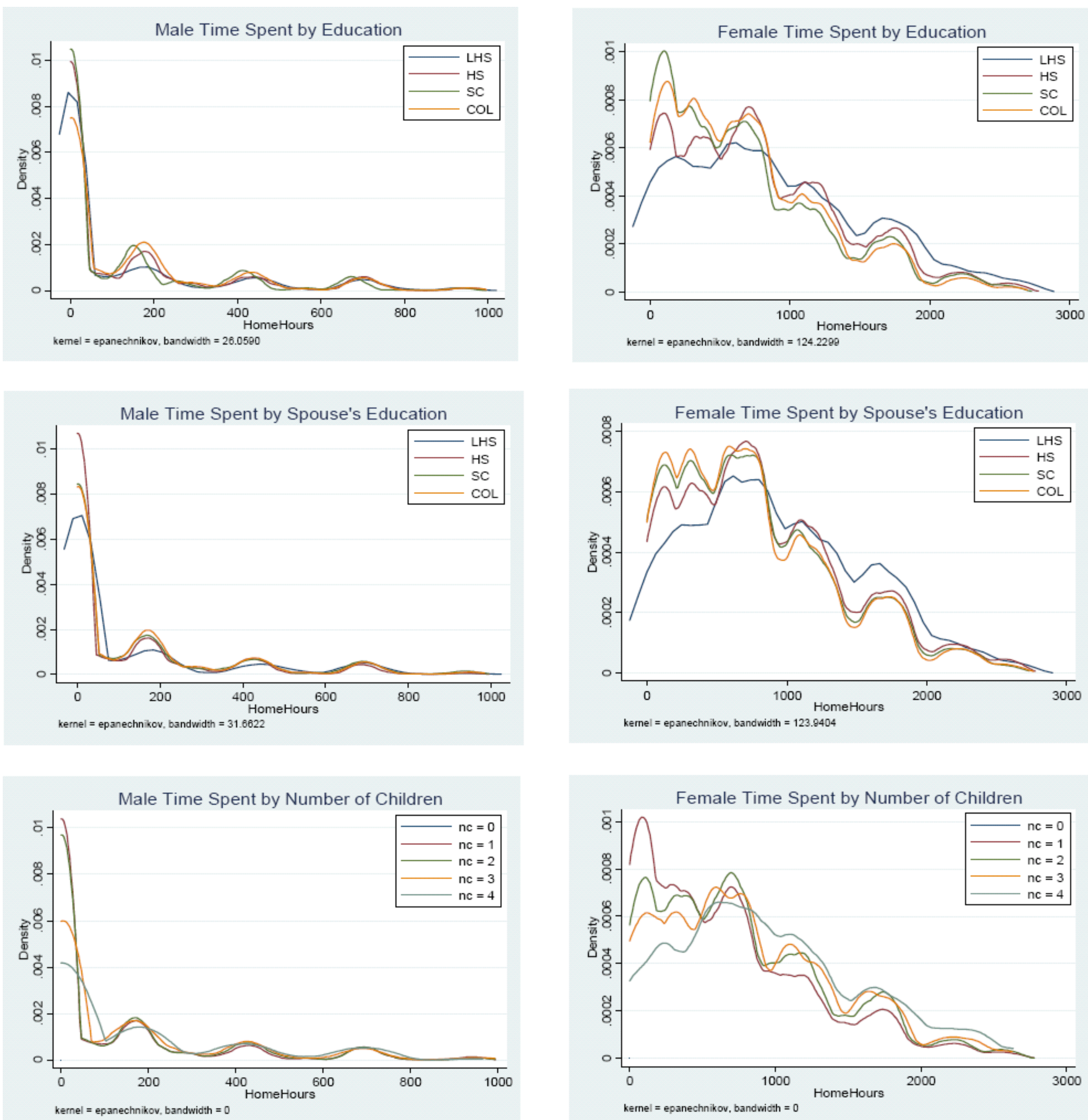
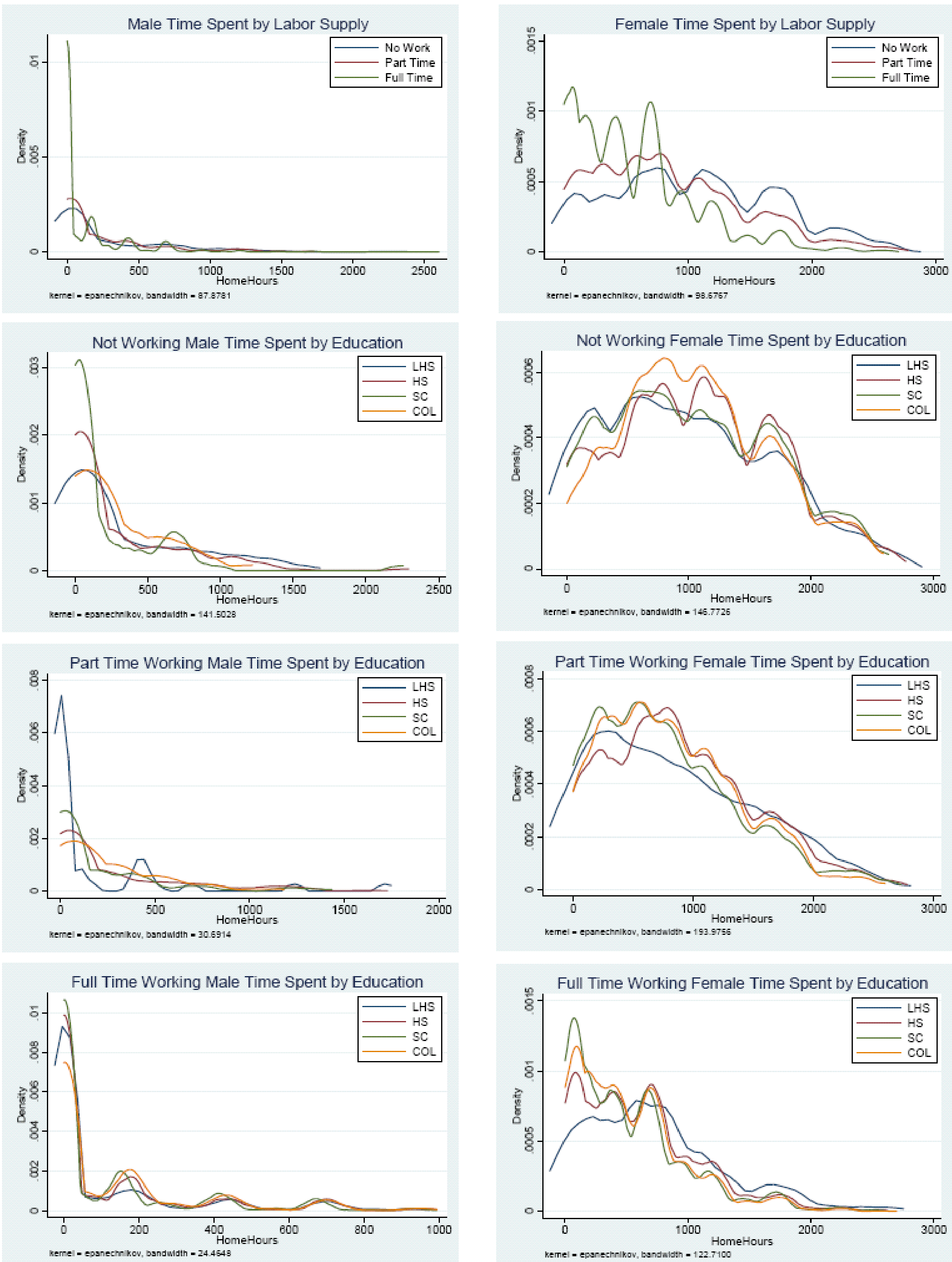


Figure 3: Parental Time Densities by Labor Supply and Education



3.2 THEORETICAL FRAMEWORK

The theoretical framework builds on Becker and Barro (1989)(9) and the literature which generalizes it (see Alvarez (1999)(5), Doepke (2005)(28) among others). Our model is a dynastic model with altruistic preferences in which each individual in a generation makes consumption, fertility, time spend with children and labor supply decisions sequentially over the life cycle. Households may consist of individuals or a couple making decisions. We model couples decisions as a noncooperative game and solve for a Markov Perfect Equilibrium (MPE) in pure strategies. We do not model household formation and dissolution as choices; instead, marriage and divorce and assumed to evolve stochastically, but the process depends on the individual and household time invariant as well as endogenous characteristics (such as number of children, human capital accumulated with experience etc.). Individuals therefore, take into account the effect of choices on probability of marriage and divorce, thus these variables are endogenous in a predetermined sense.

There are two types of individuals, female and male denoted by $\sigma = f, m$, respectively. Adults live for T periods in which they make decisions, $t \in \{0, 1, .., T\}$. An adult from generation $g \in \{0, ...\infty\}$ makes choices of consumption $c_{\sigma t}$, and discrete labor supply decision $h_{\sigma t} \in \pi_h$ (e.g. not work, part time, full time), time spent with children $d_{\sigma t} \in \pi_h$ and a birth decision $b_t \in \{0, 1\}$. We assume that only females make the birth decision, thus we omit the gender subscript. The gender dummy of a child born in period t is denoted by $I_{\sigma t}$, it takes the value 1 if the child is of gender σ and 0 otherwise. We denote the vector of labor supply choice in period t by $H_{\sigma t} = \{h_{\sigma 0,..}, h_{\sigma t-1}\}$, to capture the labor market experience of the individual at the beginning of the period. We denote by $N_{\sigma t}$ the total number of children at the beginning of period t. We assume that if there is a birth in the household in period t the child belongs to both spouses in the household, however, since individuals may divorce and remarry or have children when single (female only), the number of children of each spouse in the household may be different. $D_{\sigma t} = \{d_{\sigma 0,..}, d_{\sigma t-1}\}$ is a vector of time invested in each of spouse own children up to period t. An individual time invariant characteristics are denoted by x_σ; it includes variables such as education, race and a skill. We denote the spouse of an individual by $-\sigma$, thus $x_{-\sigma}$ is the spouse's characteristics, if the individual is married.

The vector $x_{\sigma t}$ denotes the persistent state variables at the beginning of period t; it includes x_σ, $N_{\sigma t}, H_{\sigma t}, D_{\sigma t}$ as well as the gender dummies of each child $(I_{\sigma 0 \ldots} I_{\sigma t})$ and the total time invested in each child by the other parent (if the child's parent is the current spouse it is $D_{-\sigma t}$).

The *time invariant* state variables of a child of spouse σ is denoted by x'_σ; the production function of the child's characteristics is a stochastic function which depends on the parents' total input of time over the life cycle, D_s, where s indexes the child's year of birth. Denote the stochastic outcome function of a child born in period s by $m(x'_\sigma | x_f, x_m, D_s)$.

The stochastic time invariant state variables of the child also depend on the parent's time invariant traits such as education and skill level. Although we do not model explicitly the marriage decisions, marriage outcomes depend stochastically on the individual characteristics; thus the child's spouse characteristics depend stochastically on the child's characteristics: $G(x'_{-\sigma} | x'_\sigma)$.

We assume that the earnings of individuals depend on their time invariant characteristic, such as education and a given skill endowment, the human capital accumulated with experience of working full time and part time in the past, and current level of labor supply. The earnings function in periods t is given by $w_{\sigma t}(x_\sigma, H_{\sigma t-1}, h_{\sigma t})$. Earnings of individuals with the same productive characteristics depend on their other time invariant characteristics such as gender and race capturing labor market discrimination.

Assume that each period there are preference shocks to the utility associated with each choice, denoted by $\varepsilon_{\sigma t} = [\varepsilon_{\sigma 1t}, .., \varepsilon_{\sigma t K_\sigma}]$; the shocks $\varepsilon_{\sigma kt}$ are drawn independently across choices, periods, individuals and generations from a distribution function F_ε. The shocks are also conditionally independent (of all state variables). The individual per period utility depends on their current earnings and their spouse's current earning, leisure, whether there is a birth in that period and the preference shock $\varepsilon_{\sigma kt}$. The discount factor of the valuation of the children's utility is given by $\lambda N_\sigma^{1-\nu}$, where N_σ is the total number of children a person has at the end of the life cycle. β is the annual discount factor. Denote by $U_{\sigma g}$ the discounted expected lifetime utility of an individual in generation g at period 0

$$U_{\sigma g} = E_0(\sum_{t=0}^{T} \beta^t \left[u(w_{\sigma t}, w_{-\sigma t}, x_\sigma, b_t, h_{\sigma t}, d_{\sigma t,} N_{\sigma t}) + \varepsilon_{\sigma k t}\right] + \beta^T \lambda \frac{N_\sigma^{1-\nu}}{N_\sigma} \sum_{t=0}^{T} b_t \left(\sum_{\sigma} I_{\sigma t} U_{\sigma g+1} \right)) \tag{3.1}$$

The first element on the right hand side is the per period utility of an adult in generation g of gender σ. We do not formally model bargaining and allocation of consumption within the households, and assume that the per period utility from consumption depends on the current earning, the spouse's current earnings and number of children; our formulation is consistent with no borrowing or saving and transfers between spouses. Specifically, the consumption of spouses depend on their own labor market income and labor supply, their spouses labor supply and income and on the number of children. Alternatively, if the utility in separable and linear in consumption, the formulation is consistent with wealth maximization and transfers between spouses (in addition to utility from leisure and children). We further discuss the functional form assumptions in Section 4. The per-period utility also depends on whether there is a birth in the household capturing costs of birth, the number of children (which captures the reduction in consumption due to the costs of raising children and possibly a utility value of having the children) and leisure. Because the labor supply and time spent with children choices are discrete, the current level of leisure is fully captured by $h_{\sigma t}, d_{\sigma t}$. The second element is the altruistic component of the preferences; it captures the average expected lifetime utility of a child weighted by the discount, $\lambda N_\sigma^{1-\nu}$, which is assumed to be concave in the number of children, thus $0 < \nu < 1$. Our formulation captures several differences between men and women, therefore, the expected utility of a child depends on the child's gender. There per-period utility of females and males may differ when there is birth, and labor market earnings of males and females with the same level of skills, education and experience may differ due to discrimination, which we assume to be exogenous. Furthermore, utility from own earnings and the spouse's earnings, may differ by gender, capturing differences in allocation of consumption within households.

Let $x_t = (x_{ft}, x_{mt})$ denote the persistent state variables of the spouses in the household and $\varepsilon_t = (\varepsilon_{ft,}\varepsilon_{mt})$ the vectors of preference shocks of both spouses. Denote specific choices made in each period by $k_{\sigma jt}$ and the spouse's choices are denoted by $k_{-\sigma it}$.The vector of choices made by both spouses in the household in period t is denoted by $k_{jit} = (k_{\sigma jt}, k_{-\sigma it})$ with j denoting the choices of individual σ and i denoting the choices of their spouse $-\sigma$. Also denote by $F(x_{t+1}|x_t, k_{jit})$ the stochastic transition function of the state variables, conditional on last period household state variables and choices. We assume that all transition functions are known to all individuals in all periods and generations. At the beginning of the period, all the household state variables are common knowledge, including the individual taste shocks.

We assume that each period decisions are made in two stages. In the first stage labor supply,investment, transfers to children are chosen by each individual, and birth decisions by the female simultaneously. In a second stage consumption allocation is made. In a second stage consumption allocation is made according to the sharing rules.

A Markov strategy profile for σ in the game is a vector $k_\sigma = [k_{\sigma 0}(x_t, \varepsilon_t), ., k_{\sigma T}(x_T, \varepsilon_T)]$, which describes the action for all possible household states variables x_t, ε_t in every period, where $k_{ft}(x_t, \varepsilon_t) = (d_{ft}(x_t, \varepsilon_t), h_{ft}(x_t, \varepsilon_t), b_t(x_t, \varepsilon_t))$ and $k_{mt}(x_t, \varepsilon_t) = (d_{mt}(x_t, \varepsilon_t), h_{mt}(x_t, \varepsilon_t))$ are the period t decisions in every state . Note that $k_{\sigma t}(x_t, \varepsilon_t)$ is a mapping from all possible states to K_σ possible combination of choices every period: $k_0, .., k_{K_\sigma}$. Let $k_t = (k_{\sigma t}(x_t, \varepsilon_t), k_{-\sigma t}(x_t, \varepsilon_t))$ denote an element t in a specific strategy profile of both spouses. The strategy profile maps the state variables into choices of both spouses, where a specific set of choices $k_{jit} = (k_{\sigma jt}, k_{-\sigma it})$.

Under the assumption of stationarity, we omit the generation index g. We first define the ex-ante value function V_σ as the discounted sum of future utilities. This is the the discounted sum of future utilities for household member σ before individual-specific preference shocks are observed and actions taken. Lets also define by $p(k_t|x_t)$ the conditional ex ante (again before ε_t is observed) probability that household action profile k_t will be chosen conditional on state x_t. For $t < T$ the ex ante value function can therefore be written as

$$\begin{aligned} V_\sigma(x_t) &= \sum_{k_t} p(k_t = s|x_t) \left[u(k, x_{\sigma t}) + \beta \sum_{x_{t+1}} V_\sigma(x_{t+1}) F(x_{t+1}|x_t, k_t) \right] \\ + \sum_{s=1}^{K_t} E_\varepsilon[\varepsilon_{\sigma t}|k_t &= s] p(k_t = s|x_t) \end{aligned} \tag{3.2}$$

where E_ε denotes the expectation operator with respect to the individual-specific preference shocks.

Let $\upsilon_\sigma(k_{jit}; x_t)$ denote individual σ's best response continuation value net of the preference shocks playing strategy $k_{\sigma jt}$ conditional on the spouse playing strategy $k_{-\sigma it}$. This can be written as:

$$\upsilon_\sigma(k_{jit}; x_t) = u(k_{jit}, x_{\sigma t}) + \beta \sum_{x_{t+1}} V_\sigma(x_{t+1}) F(x_{t+1}|x_t, k_{jit}). \tag{3.3}$$

Recall that a vector of choices for a household is given by $k_{jit} = (k_{\sigma jt}, k_{-\sigma it})$. Thus, given a spouse strategy $k_{-\sigma it}$ a vector of choice $k_{\sigma jt}$ is optimal if $\upsilon_\sigma(k_{\sigma jt}, k_{-\sigma it}; x_t) + \varepsilon_{\sigma jt} \geq \upsilon(k_{\sigma j't}, k_{-\sigma it}; x_t) + \varepsilon_{\sigma j't}$ for $k_{\sigma j't}$. Thus, we can characterize the probability distribution over $k_{\sigma jt}$ for all j and write the conditional ex ante choice probabilities of the choice profile given a spouse's strategy profile:

$$p_{\sigma jt}(k_{\sigma jt}|k_{-\sigma it}, x_t) = \int \left[\prod_{k_{\sigma jt} \neq k_{j'it}} 1\{\upsilon_\sigma(k_{jit}; x_t) - \upsilon_\sigma(k_{j'it}; x_t) \geq \varepsilon_{\sigma jt} - \varepsilon_{\sigma j't}\} \right] dF_\varepsilon \tag{3.4}$$

where $\upsilon_\sigma(k_{jit}; x_t) - \upsilon_\sigma(k_{0it}; x_t)$ is the differences in the ex-ante conditional valuation when individual σ chooses $k_{\sigma jt}$ and the valuations when $k_{\sigma j't}$ is chosen *given* that the spouse chooses $k_{-\sigma it}$. Notice that the choices $k_{\sigma jt}$ and $k_{\sigma j't}$ are chosen according to the strategy k_σ which maps for every period state variables (x_t, ε_t) into choices, and given a spouse choices, we describe the probability distribution over the choices of an individual when the strategy is optimal. Because the conditional independence of the shocks, the household strategies probabilities are given by

$$p(k_t|x_{t+1}) = p_{\sigma jt}(k_{\sigma jt}|k_{-\sigma it}, x_t) \times p_{-\sigma it}(k_{-\sigma it}|x_t). \tag{3.5}$$

Define the intergenerational transition function of the persistent state variables of a child born in period s in the parent's life cycle by

$$M(x_0'|x_f, x_m, D_s) \equiv m(x_\sigma'|x_f, x_m, D_s) G(x_{-\sigma}'|x_\sigma').$$

This function captures the stochastic outcomes of the child in terms of the child time invariant characteristics and the child's spouse characteristics, given the parents' time invariant characteristics and time investment in the child. The ex-ante conditional best response function net of the preference shock in the final period of the life cycle T is given by

$$\upsilon_\sigma(k_{jiT};x_T) = u(k_{jiT},x_{\sigma T}) + \beta\lambda\frac{(N_{\sigma T}+b_T)^{1-\upsilon}}{(N_{\sigma T}+b_T)}\overline{V}_{N\sigma}(k_{jiT};x_T) \tag{3.6}$$

Where $\overline{V}_N(x_T)$ is sum of the expected valuation over all children born up to period T plus the valuation of a child born in period T if there is birth

$$\begin{aligned}\overline{V}_N(k_{jiT};x_T) \equiv & \sum_{s=0}^{T-1}\left[b_s\sum_\sigma I_{\sigma s}\sum_{x_0'}V_{\sigma s}(x_0')M(x_0'|x_f,x_m,D_s)\right] \\ & +b_T\sum_\sigma p_\sigma\sum_{x_0'}V_{\sigma T}(x_0')M(x_0'|x_f,x_m,D_T)\end{aligned} \tag{3.7}$$

Note that D_T and D_s for $s<T$ are both functions of k_{jiT}. In the final period of the life cycle, the valuation function (Equation 3.6) depends on current utility, and the discounted expected value of the children's valuation functions. The first element of Equation 3.7 is the expected valuation of the existing children at the beginning of period T, which state variables depend on past parental time input and the current period inputs. The second element is the expected value of a child born in period T for which the gender is unknown at the beginning of the period. Thus, this element depends on the birth decision and parental time input. We assume that all children become adults after period T and their state variables are unknown until then regardless of the time of birth.

We solve for a Markov Perfect Equilibrium of the game; restricting attention to pure strategies and do not consider mixed strategies.

Definition 1 (Markov perfect equilbrium). *A strategy profile k° is said to be a Markov perfect equilibrium if for any $t\leq T$,$\sigma\in\{m.f\}$, and $(x_t,\varepsilon_t)\in(X,R^{K_f+K_m})$,*

1. $\upsilon_\sigma(k^\circ_{jit};x_t)+\varepsilon_{\sigma jt}\geq\upsilon_\sigma(k^0_{j'iT};x_t)+\varepsilon_{\sigma j't}$;
2. *all players use Markovian Strategies*

In general a pure strategy Markovian perfect equilibrium for complete information stochastic games may not exist, however, we imposed sufficient conditions on the primitives of our game and show that there exist at least one pure strategies Markov perfect equilibrium. To show this results, we use some of the properties and definitions of supermodular games on lattice theory (see Milgrom and Roberts(1990)(59), Milgrom and Shannon (1994)(60), and Tokis(1998)(82) for examples these properties). A binary relation $\geq$ on a non-empty set is a partial order if it is reflexive, transitive, and anti-symmetric. A partially ordered set is said to be a lattice if for any two elements the supremum and infimum are elements of the set. A 2 person game is said to be supermodular if the set of actions for each player σ is a compact lattice and the payoff function is supermodular in k_σ for fixed $k_{-\sigma}$ and satisfies increasing differences in $(k_\sigma, k_{-\sigma})$. Following Watanabe and Yamashita (2010)(83), if the continuation values in every period and state satisfy the conditions below, the game is supermodular and there exists a pure strategies Markov perfect equilibrium. Following the convention, we use $\vee$ to denote the supremum of two elements and $\wedge$ to denote the infimum of two elements.

Condition 1 (S). *$\upsilon_\sigma(k_{\sigma t}, k_{-\sigma t}, x_t)$ is supermodular in $k_{\sigma t}$ for any $x_{\sigma t}$ and $k_{-\sigma t}$ if*

$$\upsilon_\sigma(k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + \upsilon_\sigma(k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \geq \upsilon_\sigma(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + \upsilon_\sigma(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \quad (3.8)$$

for all $(k'_{\sigma t}, k_{\sigma t})$.

Condition 2 (ID). *$\upsilon_\sigma(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$ has increasing differences in $(k_\sigma, k_{-\sigma})$ for any $x_{\sigma t}$ if*

$$\upsilon_\sigma(k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) - \upsilon_\sigma(k_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) \geq \upsilon_\sigma(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) - \upsilon_\sigma(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \quad (3.9)$$

for all $k'_{\sigma t} \geq k_{\sigma t}$ and $k'_{-\sigma t} \geq k_{-\sigma t}$.

Watanabe and Yamashita (2010)(83) provide sufficient conditions on the stochastic transitions functions and the per period utility for the these exist a pure strategy Markov perfect equilibrium. These conditions impose restrictions on the functional forms of the per period utility sharing rules, wage functions, value of kids, and the return investment in children. In the implementation section we discuss these restrictions further once the functional of these primitives are specified and provide a proof.

3.3 ESTIMATION

We use a representation of the valuation function in terms of the model's primitives and choice probabilities which allows for the estimation of the problem in several steps (see Gayle, Golan, Soytas (2011)(35), details on the estimation of discrete choice dynastic models). The estimator accommodates the multiple equilibria issue. The difficulty of estimating the model is due to the non-standard nature of the problem. While the problem can be solved with a nested fixed point algorithm, it becomes computational intensive quickly, limiting the scope of the problem that can be analyzed. The alternative representation developed of the continuation value of the intergenerational problem enables us to derive the necessary representation and apply the Hotz and Miller (1993)(43)estimation technique for single agent problems to the dynastic problem. We use the estimator developed in a companion paper Gayle, Golan, Soytas (2011)(35),, beginning with the following representation of the problem,

$$\begin{aligned} \upsilon_{\sigma}(k_{jit};x_t) &= u_{\sigma}(k_{jit},x_t) \\ &+\sum_{s=t+1}^{T}\beta^{s-t}\sum_{x_s}\left\{\left(\sum_{k_s}[u_{\sigma}(k_s,x_s)+E_{\varepsilon}(\varepsilon_{\sigma s}|k_s=s)]p(k_s=s|x_s)\right)F(x_s|x_t,k_{jit})\right\} \\ &+\lambda\beta^{T-t}\sum_{x_0}V(x_0')H(x_0'|x_t,k_{jit}) \end{aligned} \tag{3.10}$$

where $F(x_s|x_t,k_{jit})$ is the $s-t$ transitions, $H(x_0'|x_t,k_{jit})$ is weighted generation transitions, and $V(x_0)$ $(=[V_f(x_0),V_m(x_0)]')$ is a vector of the ex-ante . The transition function $H(x_0'|x_t,k_{jit})$ can write as recursive function of $F(x_{t+1}|x_t,k_{jit})$, $M(x_0'|x_f,x_m,D_s)$, $N_{\sigma T}$, b_s, p_σ and $1-\nu$. Define the ex-ante conditional lifetime utility as period t, exclusion the dynastic component as:

$$\begin{aligned} U_{\sigma}(k_{jit},x_t) &= u_{\sigma}(k_{jit},x_t) \\ &+\sum_{s=t+1}^{T}\beta^{s-t}\sum_{x_s}\left\{\left(\sum_{k_s}[u_{\sigma}(k_s,x_s)+E_{\varepsilon}(\varepsilon_{\sigma s}|k_s=s)]p(k_s=s|x_s)\right)F(x_s|x_t,k_{jit})\right\} \end{aligned}$$

Therefore we can write an alternative representation for the ex-ante value function as time t :

$$\begin{aligned} V_{\sigma}(x_t) &= \sum_{k_{-\sigma it}}\left\{p(k_{-\sigma it}|x_t)\sum_{k_{\sigma jt}}\left[U_{\sigma}(k_{jit},x_t)+E_{\varepsilon}(\varepsilon_{_{\sigma jt}}|k_{jit},x_t)\right]p_t(k_{\sigma jit}|x_t)\right\} \\ &\quad+\sum_{k_{-\sigma it}}\left\{p(k_{-\sigma it}|x_t)\sum_{k_{\sigma jt}}\left[\lambda\beta^{T-t}\sum_{x_0}V(x_0)H(x_0|x_t,k_{jit})\right]p_t(k_{\sigma jit}|x_t)\right\} \end{aligned} \tag{3.11}$$

Equation (3.11) is satisfied at every state vector x_t, and since the problem is stationarity over generation at period 0 we express it as a matrix equation:

$$\begin{aligned} V(X_0) &= P(X_0)U(X_0)+e(X_0,P(X_0))+\lambda\beta^T P(X_0)H(X_0)V(X_0) \\ &= [I_{2S(X)}-\lambda\beta^T P(X_0)H(X_0)]^{-1}[P(X_0)U(X_0)+e(X_0,P(X_0))] \end{aligned} \tag{3.12}$$

The terms on the right hand side of Equation 3.12 are the intergeneration and the per period discount factors, the household choice probability matrix, the intergeneration state transition matrix, the ex-ante conditional lifetime utility, and the expected purveyances shocks. In matrix notation $V(X_0)=[V(x_0)]_{x_0\in X_0}$ is $2S(X_0)\times 1$ vector of expected discounted sum of future utility; $P(X_0)$ is $2S(X_0)\times(S(K)\cdot 2S(X_0))$ dimensional matrix consisting of the household choice probability $p(k|x_0)$ in rows x_0 and $S(X)$+x_0 and columns (k,x_0) and (k,$S(X)$+x_0), zeros in rows x_0 and $S(X)$+x_0 and columns (k,x_0') and (k,$S(X)$+x_0') with $x_0'\neq x_0$; $e(X_0,P(X_0))$is the $2S(X_0)\times 1$ vector of expected preference shocks with element $[\sum_{k_{-fi}}E_{\varepsilon}(\varepsilon_{_{fj}}|k_{ji},x)p(k_{fji}|x)p(k_{-fi}|x),\sum_{k_{-mi}}E_{\varepsilon}(\varepsilon_{_{mj}}|k_{ji},x)p(k_{mji}|x_t)p(k_{-mi}|x_t)]'_{x\in X_0}$; and $I_{2S(X)}$ denotes the $2S(X_0)$-dimensional identity matrix. The second line in Equation (3.12) is a direct implication of the dominant diagonal property, which implies that the matrix $[I_{2S(X)}-\lambda\beta^T P(X_0)H(X_0)]$ is invertible.

Under the assumption that $\varepsilon_{_{\sigma s}}$ is distributed i.i.d. type I extreme value then Hotz and Miller (1993)(43), inversion implies that

$$\begin{aligned} \log\left(\frac{p_{\sigma jt}(k_{\sigma jt}|k_{-\sigma it},x_t)}{p_{\sigma jt}(k_{\sigma 0t}|k_{-\sigma it},x_t)}\right) &= \\ & U_{\sigma}(k_{jit},x_t)-U_{\sigma}(k_{0it},x_t) \\ & +\lambda\beta^T\sum_{x_0}V(x_0)[H(x_0|x_t,k_{jit})-H(x_0|x_t,k_{0it})] \end{aligned} \tag{3.13}$$

for $\sigma \in \{f, m\}, k_{jit} \neq k_{0it}$. Using equation (3.13) we then use a simulated method of moment estimation techniques developed in Hotz, Miller, Sanders and Smith (1994)(44). In the first step we estimate the transition functions and conditional best response probabilities from the data. Starting at age seventeen we use the estimate in the first step to simulate lifetime paths for each value of the state space. Using the formulate in equation (3.12), we compute and estimate of $V(X_0)$ from the simulated data. Similarly we simulated paths for each value of the state space at age greater seventeen which to obtain and estimate of the for Next we simulate of $U_\sigma(k_{jit}, x_t)$. Using the estimates of the conditional best response probabilities, transition functions, $V(X_0)$, and $U_\sigma(k_{jit}, x_t)$, we form an empirical counterpart to equation (3.13) and estimate the parameters of our model using a 2-step GMM estimator.

3.3.1 Empirical Implementation

We describe the choice set specifications, functional forms of model which we estimate and discuss existence and implications.

3.3.1.1 Choice sets We set the number of periods in each generation $T = 39$ and measure the individual's age where $t = 0$ is age 17. Below we summarize the decision process of males and females for possible choice combinations. Define an indicator variable $\mathbb{I}_{k_{\sigma t}}$ where $\mathbb{I}_{k_{\sigma t}} = 1$ if the action $k_{\sigma t}$ is chosen and $\mathbb{I}_{k_t \sigma} = 0$ otherwise. Females have 16 mutually exclusive choices each includes a level of labor market time, time spent with children and a birth decision. Thus, with 3 levels of labor supply corresponding to no work, part time work, and full time work (i.e. $h_{ft} \in \{0, 1, 2\}$). These levels are defined using the 40 hours week; an individual working less three hours per week is classified as not working, individuals working between 3 and 20 hours per week are classified as working part time, while individuals working more than 20 hours per week are classified as working full time. There are 3 levels of parental time with kids corresponding to no time, low time, and high time. To control for the fact female spends significantly more time with kids than male we used a gender specific categorization. We used the 50th percentile of the distribution of parental time with kids as the threshold for low versus high parental time with children, thus a parent spending parent spending greater than zero but less than the 50th percentile is classified as spending low time with kids and greater or equal to the 50 percentile is classified as spend high time with kids (i.e. $d_{\sigma t} \in \{0, 1, 2\}$). Finally, birth is a binary variable equal one of the mother give birth child in that year and zero otherwise (i.e. $b_t \in \{0, 1\}$). Table 5 presents the summary of these 16 mutually exclusive choices.

Males have 9 mutually exclusive choices since they do not have a birth decision; there labor market and parental time decisions defined the same way as female except that the parental time threshold is defined using the male distribution of parental time hours. The second panel in Table 3 presents the summary of the males choice set. Let sets $\mathcal{H}_{P\sigma}$ and $\mathcal{H}_{F\sigma}$ index the choices that involve working part time and full time in the labor market respectively and let $\mathcal{H}_\sigma$ be the choice set for each gender σ.

Individual utility is a function of consumption, leisure and number of children which affects consumption. The per period utility of an individual is composed of two parts; utility from own and spouse's current income and number of children and the utility from leisure. We assume the following functional forms for the utility from income for a married (or for cohabitation) individual in period t

$$u_{1\sigma t} = \alpha_\sigma w_{\sigma t} \sum_{k_{t-s} \in \mathcal{H}_{F\sigma} \cup \mathcal{H}_{P\sigma}} \mathbb{I}_{\sigma k_{t-s}} + \alpha'_\sigma w_{-\sigma t} \sum_{k_{t-s} \in \mathcal{H}_{F-\sigma} \cup \mathcal{H}_{P-\sigma}} \mathbb{I}_{-\sigma k_{t-s},} + \alpha_{\sigma N}(N_t^{17} + b_t) \quad (3.14)$$

where N_t^{17} is the effective number of children less than 17 years old. The per-period utility from income for a single individual is

$$u_{1\sigma t} = \alpha_\sigma w_{\sigma t} \sum_{k_{t-s} \in \mathcal{H}_{F\sigma} \cup \mathcal{H}_{P\sigma}} \mathbb{I}_{\sigma k_{t-s}} + \alpha_{\sigma N}(N_t^{17} + b_t) \quad (3.15)$$

This formulation is consistent with each spouse consuming a share of their income net of their share of costs of children and a transfer from the spouse. Assuming no borrowing and saving, one can restrict the coefficients on the income, spouse's income and number of children so that the total value of consumption equals the total household income net of costs of children and the per-period budget constraint is satisfied. However, since we do not have data on consumption or costs of children, the coefficients on the number of children also captures non-pecuniary utility from children and cannot be identified separately from the monetary costs of raising children.

Table 5: Discrete Choice Set of Structural Model

	Decisions		
Choice	Labor Market Work	Child Birth	Child Care Hours
	Female		
1	None	None	None
2	Part time	None	None
3	Full Time	None	None
4	Full Time	Yes	None
5	None	None	Low
6	Part Time	None	Low
7	Full Time	None	Low
8	None	Yes	Low
9	Part Time	Yes	Low
10	Full Time	Yes	Low
11	None	None	High
12	Part Time	None	High
13	Full Time	None	High
14	None	Yes	High
15	Part Time	Yes	High
16	Full Time	Yes	High
	Male		
1	None	NA	None
2	Part Time	NA	None
3	Full Time	NA	None
4	None	NA	Low
5	Part Time	NA	Low
6	Full Time	NA	Low
7	None	NA	High
8	Part Time	NA	High
9	Full Time	NA	High

We assume that the preferences are additive in consumption and leisure. We there defined the per period disutility from working for each gender as

$$u_{2\sigma t} = \sum_{k_t \in \mathcal{H}_\sigma} \theta_{\sigma k_t} \mathbb{I}_{k\sigma t} \tag{3.16}$$

where $\theta_{\sigma k_t}$ are the coefficients associated with each choice, thus capturing the disutility from any combination of time spent with children and at work, thus capturing the value of leisure. For females, the disutility from working and spending time with children also depends on whether there is a birth or not in that period, whereas for males, the only effect of birth is through the effect of an additional child α_{mN} . For notational ease we omit age, education, and race but all the above utility parameters are allowed to vary by these characteristics.

3.3.1.2 Labor Market Earnings Individual's earnings depend on his/her characteristics, $x_{\sigma t}$. Let $z_{\sigma t}$, be a subset of $x_{\sigma t}$, which includes age, age squared and Ed_σ, an education dummy variables indicating whether the individual has high school, some college or college (or more) education interacted with age respectively[2]. Let η_σ be the individual specific ability which is assumed to be correlated with the individual specific time invariant observed characteristics.. Earnings are assumed to be the marginal productivity of workers, and is assumed to be exogenous, linear additive and separable across individuals in the economy. The earnings equations for female and male are given by:

$$w_{\sigma t} = \exp(\delta_{0\sigma} z_{\sigma t} + \sum_{s=0}^{\rho} \delta^{pt}_{\sigma,s} \sum_{k_{t-s} \in \mathcal{H}_{P\sigma}} \mathbb{I}_{k_{t-s}\sigma} + \sum_{s=1}^{\rho} \delta^{ft}_{\sigma,s} \sum_{k_{t-s} \in \mathcal{H}_{Fm}} \mathbb{I}_{k_{t-s}\sigma} + \eta_\sigma) \tag{3.17}$$

where the earnings equation depends on experience accumulated while working part time and full time, and the current level of labor supply. We assume $\rho = 4$, and the depreciation and different values of human capital accumulated while working part-time and full time as well as the depreciation rates are captured by $\delta^{pt}_{\sigma,s}$ and $\delta^{ft}_{\sigma,s}$, respectively.

[2] Level of education Ed_σ is a discrete random variable in the model where it can take 4 different values for: less than high school (LHS), high school (HS), some college (SC) and college (COL).

3.3.1.3 Production Function of Children

Parental time investment in children affect the future educational outcome of the child which is denoted by Ed'_σ. and innate ability η'_σ, both affecting the child's earnings (see Equation 3.17).

The state vector for the child in the first period of her life cycle $x'_{0\sigma}$ is determined by the intergenerational state transition function $M(x'_0|x_f, x_m, D_s)$ specifically,

$$M(x'_0|x_f, x_m, D_s) = \Pr(\eta'_\sigma \mid Ed'_\sigma) \Pr(Ed'_\sigma \mid x_f, x_m, D_s) \Pr(\eta'_\sigma \mid Ed'_\sigma) \Pr(Ed'_{-\sigma 0} \mid Ed'_\sigma) \tag{3.18}$$

Thus, we assume that the parental inputs and characteristics (parents education and fixed effects) determines educational outcomes according to probability distribution $\Pr(Ed'_\sigma \mid x_f, x_m, D_s)$. The state vector of inputs contains the cumulative investment variables (low time and high time) of each parent up to period T. We assign each child in the household the average time investment assuming all children in the household receive the same time input. Parents's characteristics include the education of the father and mother, their individual specific effects and race. Once the education level is determined, it is assumed that the ability η'_σ is determined according to the probability distribution $\Pr(\eta'_\sigma \mid Ed'_\sigma)$. The spouse's education is also determined after the realization of the child's education according to the distribution $\Pr(Ed'_{-\sigma 0} \mid Ed'_\sigma)$, potentially capturing assortative mating. The above form of the transition allows us to estimate the equations separately for the production function of children given as the first two probabilities, and the marriage market matching given as the last term.

3.3.1.4 Existence of MPE in Pure Strategies

We need one final assumption to guarantee that there exist a MPE in pure strategies.

Assumption 1: For an increasing levels of $\widehat{Ed_\sigma}$

$$\Pr(\widehat{Ed_\sigma}|k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) - \Pr(\widehat{Ed_\sigma}|k_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) \geq \Pr(\widehat{Ed_\sigma}|k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) - \Pr(\widehat{Ed_\sigma}|k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$$

for all $k'_{\sigma t} \geq k_{\sigma t}$ and $k'_{-\sigma t} \geq k_{-\sigma t}$.

The property implies that the differences in outcomes of children in terms of higher x'_0 are weakly higher the larger the existing stock of investment. Thus, if there are complementarities in time investment of parents or if the increase in outcomes is independent of the spouse's investment, the condition is satisfied. Table 4 shows that this condition is satisfied. It is important that we estimated the education production function outside the main estimation hence we can verify that these exist a MPE in pure strategies before the imposing it. This guarantees that our estimator is well define over the parameters space.

Proposition 1. *Under Assumption 1 and given the specification in equation (3.14), (3.16), (3.17) and (3.18); there exist a MPE in Pure Strategy.*

3.4 RESULTS

As noted in the estimation section we used a multi-stage estimation technique. As such we present the results in three stages. The first stage presents the estimates of the intergeneration education production function, the earnings equation, the unobserved skills function, the marital status transition functions, and the marriage assignment functions. All these functions are fundamental parameters of our model which are estimated outside the main estimation of the preference, discounts factors, household sharing rules (coefficient on own and spouse earnings in the utility function), and the net costs of raising children parameters. The first stage estimates also include equilibrium objects such as the conditional choice probabilities and the best response functions. The second stage presents estimates of the intergenerational and intertemporal discount factors, the preference parameters, the household sharing rules, and child care cost parameters. The third and final stage presents counter factual estimates of the return to parental time investment and the value of children.

3.4.1 First Stage Estimates

3.4.1.1 Intergenerational Education Production Function A well known problem with the estimation of production functions is the simultaneity of the inputs. As is clear from the structural model the intergenerational education production function suffers from a similar problem. However, because the output of the intergeneration education production (i.e. completed education level) is determined over generations while the inputs, such as parental time investment, are determined during the life cycle, we can treat these inputs as predetermined and use instruments from within the system to estimate the production function.

Table 6 presents results of a Three Stage Least Square estimation of the system of individual educational outcomes. The estimation uses mother's and father's labor market hours over the first 5 years of the child's life as well as linear and quadratic terms of mother's and father's age on the 5th birth day of the child as instruments. The estimation results show that a child who's mother has a college education has a significantly higher probability of graduating from college and a lower probability of only being a high school graduate, while if a child's father has some college or college education the child has a higher probability of graduating from college.

We measure parental time investment as the sum of the parental time investment over the first 5 years of the child's life. Total time investment is a variable that ranges between 0 and 10 since low parental investment is coded as 1 and high parental investment is coded as 2. The results in Table 6 shows that while mothers time investment significantly increases the probability of a child graduating from college, fathers time investment significantly increases the probability of the child graduating from high and going to college. These estimates suggest that while mothers' time investment increases the probability of a high educational outcome, fathers' time investment truncates low educational outcome. However, both parents' time investment is productive in terms of children education outcomes. It is important to note that mothers' and fathers' hours spent with children are at different margins, with mothers providing significantly more than fathers. Thus the magnitudes of the discrete levels of time investment of mothers and fathers are not directly comparable since what constitutes low and high investment differs across genders.

The results in Table 6 also show that females are more likely to enter and graduate college than males. Interestingly, controlling for parental characteristics and time investment, black children have a higher probability of graduating from college as well as a higher probability of not graduating from high school than white children.

Table 7 presents the predicted probabilities of a child's education outcomes by parents education and time investment for a white male child. This exercise illustrates the quanti-

Table 6: 3SLS System Estimation the Education Production Function

(Standard Errors in parenthesis; Excluded class is Less than High School)

Variable	High School	Some College	College
High School Father	0.008	0.023	0.155
	(0.068)	(0.104)	(0.128)
Some College Father	-0.012	0.057	0.162
	(0.047)	(0.074)	(0.086)
College Father	-0.014	0.021	0.229
	(0.071)	(0.110)	(0.135)
High School Mother	0.004	0.093	0.083
	(0.057)	(0.089)	(0.107)
Some College Mother	-0.016	0.036	-0.089
	(0.054)	(0.085)	(0.098)
College Mother	-0.122	0.03	0.222
	(0.076)	(0.116)	(0.140)
Mother's Time	-0.091	-0.048	0.299
	(0.075)	(0.114)	(0.130)
Father's Time	0.153	0.273	-0.108
	(0.069)	(0.103)	(0.131)
Mother's Labor Income	0.021	-0.014	-0.004
	(0.025)	(0.039)	(0.048)
Father's Labor Income	0.015	0.018	-0.023
	(0.010)	(0.016)	(0.020)
Female	0.034	0.158	0.110
	(0.030)	(0.045)	(0.056)
Black	-0.227	-0.236	0.324
	(0.093)	(0.141)	(0.168)
Constant	0.606	-0.416	-0.889
	(0.255)	(0.396)	(0.450)
Observations	980	980	980

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID), and include individuals surveyed between 1968 and 1997. Instruments: Mother's and father's labor market hours over the child's first 8 years of life, linear and quadratic terms of mother's and fathers age when the child was 5 years old.

Table 7: The probability of white male child's education outcome

			CHILD'S EDUCATION			
Mother's Education	Father's Education	Investment	Less than high school	High School	Some College	College Graduate
Less than high school	Less than high school	NO	0.14	0.86	0.00	0.00
High School	High School	NO	0.13	0.87	0.00	0.00
Some College	Some College	NO	0.16	0.84	0.00	0.00
College Graduate	College Graduate	NO	0.29	0.71	0.00	0.00
Less than high school	Less than high school	AVG	0.14	0.59	0.24	0.03
High School	High School	AVG	0.13	0.48	0.12	0.27
Some College	Some College	AVG	0.15	0.36	0.14	0.34
College Graduate	College Graduate	AVG	0.00	0.00	0.21	0.79
Less than high school	Less than high school	MAX	0.00	0.00	0.23	0.77
High School	High School	MAX	0.00	0.00	0.00	1.00
Some College	Some College	MAX	0.00	0.00	0.00	1.00
College Graduate	College Graduate	MAX	0.00	0.00	0.00	1.00

tative magnitude of the effect of parental time investment on education outcomes. It shows that if both parents have less than a high school education and invest no parental time over the child's first five years of life, the child has a 14% chance of not completing high school and 86% chance of graduating college. However, if both parents invest the average time observed in our sample then while the chance of not completing high school does not change, the probability of some college increases to 24% and the chance of graduating college increases to 3%. If both parents invest the maximum amount of time then the probabilities of not graduating from high school or only graduating high school are zero, the probability of some college is 23% and the probability of graduating from college is 77%. This pattern is repeated for other education groups; if both parents are college graduates but do not invest then the child has no chance of going to or graduating from college. These results suggest that there are significant returns to parental time investment and in the rest of the paper we quantify these returns.

3.4.1.2 Earnings Equation and Unobserved Traits Table 8 presents the estimates of the earnings equation and the function of unobserved (to the econometrician) individual skill. The top panel of the first column shows that the age-earnings profile is significantly steeper for higher levels of completed education; the slope of the age-log-earnings profile for a college graduate is about 3 times that of an individual with less than a high school education. However, the largest gap is due to being a college graduate; the of the age-log-earnings profile for a college graduate is about twice that of an individual with only some college. These results confirm that there are significant returns to parental time investment in kids in terms of labor market because parental investment significantly increases the likelihood of higher education outcomes which significantly increases life time labor market earnings.

Table 8: Estimates of Earnings Equation

Dependent Variable: Log of Yearly Earnings

(Standard Errors in Parenthesis)

Variable	Estimate	Variable	Estimate	Variable	Estimate
Demographic Variables				Fixed Effect	
Age Squared	-4.0e-4	Female x Full time work	-0.125	Black	-0.154
	(1.0e-5)		(0.010)		(0.009)
Age x LHS	0.037	Female x Full time work (t-1)	0.110	Female	-0.484
	(0.002)		(0.010)		(0.007)
Age x HS	0.041	Female x Full time work (t-2)	0.025	HS	0.136
	(0.001)		(0.010)		(0.005)
Age x SC	0.050	Female x Full time work (t-3)	0.010	SC	0.122
	(0.001)		(0.010)		(0.006)
Age x COL	0.096	Female x Full time work (t-4)	0.013	COL	0.044
	(0.001)		(0.010)		(0.006)
Current and Lags of Participation		Female x Part time work (t-1)	0.150	Black x HS	-0.029
Full time work	0.938		(0.010)		(0.010)
	(0.010)	Female x Part time work (t-2)	0.060	Black x SC	0.033
Full time work (t-1)	0.160		(0.010)		(0.008)
	(0.009)	Female x Part time work (t-3)	0.040	Black x COL	0.001
Full time work (t-2)	0.044		(0.010)		(0.011)
	(0.010)	Female x Part time work (t-4)	-0.002	Female x HS	-0.054
Full time work (t-3)	0.025		(0.010)		(0.008)
	(0.010)	Individual Specific Effects	Yes	Female x SC	0.049
Full time work (t-4)	0.040				(0.006)
	(0.010)			Female x COL	0.038
Part time work (t-1)	-0.087				(0.007)
	(0.010)			Constant	0.167
Part time work (t-2)	-0.077				(0.005)
	(0.010)				
Part time work (t-3)	-0.070				
	(0.010)				
Part time work (t-4)	-0.010	Hausman Statistics	2296		
	(0.010)	Hausman P-Value	0.000		

Table 8 (cont'd): Estimates of Earnings Equation

N	134,007	
Number of Individuals	14,018	
R-squared	0.44	0.278

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID), and include individuals surveyed between 1968 and 1997. Yearly earnings is measured in 2005 dollars. LHS is a dummy variable indicating that the individual has completed education of less than high school; HS is a dummy variable indicating that the individual has completed education of high school but college; SC is a dummy variable indicating that the individual has completed education of greater than high school but is not a college graduate; COL is a dummy variable indicating that the individual has completed education of at least a college graduate.

The bottom panel of the first column and the second of column of Table 8 show that full time workers earn 2.6 times more than part time workers for males, and 2.3 times more than part time workers for females. It also shows that there are significant returns to past full time employment for both genders; however, females have higher returns to full time labor market experience than males. The same is not true for part time labor market experience; males' earnings are lower if they work part time in the past while the there are positive returns to the most recent female part time experience. However, part time experiences 2 and 3 years in the past are associated with lower earnings for females, these rates of reduction in earnings are however lower than that of males. These results are similar to those find in Gayle and Golan (forthcoming)(33) and maybe reflect some form of statistical discrimination in the labor market in which past labor market history reflect beliefs of employers on workers labor market attachment in the presence of hiring costs.[3] These results imply that there are significant costs in the labor market in terms of loss of human capital from spending time with kids, if spending more time with kids comes at the expense of working more in the labor market. This cost may be smaller for female than males because part time work reduces compensation less for females than males. If a female works part time for 3 years, for example, in order to invest time in kids she loses significantly less human capital than a male working part time for 3 years instead of full time. This may give rise to females

[3] These results are also consistent with part time jobs being more diffferent than full time jobs, for males more than for females.

specializing in child care; this specialization comes from the labor market and production function of child's outcome as is the current wisdom.

The unobserved skill (to the econometrician) is assumed to be a parameter function of the strictly exogenous time-invariant components of the individual variables. This assumption is used in other papers such as Macurdy (1981)(55) Chamberlain (1986)(18), Nijman and Verbeek (1992)(67), Zabel (1992)(86), Newey (1994)(66), Altug and Miller (1998)(4), and Gayle and Viauroux (2007)(32). It allows us to introduce unobserved heterogeneity to the model but at the same time maintain the assumption on the discreteness of the state space of the dynamic programming problem needed for the estimation of the structural parameters from the dynastic model. The Hausman statistic shows that we cannot reject this correlated fixed effect specification. Column 3 of Table 8 presents the estimate of the skill as function of unobserved characteristics; it shows that blacks and females have lower unobserved skill than whites and males. This could capture labor market discrimination. Education increases the level of the skill but it increases at a decreasing rate in the level of completed education. The rate of increase for blacks and females with some college and a college degree are higher than their white and male counterparts. This pattern is reversed for blacks and females with a high school diploma. Notice that the skill is another transmission mechanism through which parental time investment affects labor market earnings in addition to education.

3.4.1.3 Married Transitions and Assignment Table 9 presents the logit coefficient estimates of the one period transition from single to marriage. It shows that blacks of both genders are less likely to be married next period if they are currently single. The level of education does not have any effect on the male's transition from single to married. However a single female with a high school education is more likely to transition to marriage next period than any other level of education, while a single female with a college degree is less likely to transition to marriage next period than any other education group. This result may mean that while college education for females is valuable in the labor market it may not be as valuable in the marriage market, however, another option is that college education implies a better outside options and a higher value of being single.

Table 9 also shows the single to married transition probabilities are concave in age for both genders. The number of children, while not affecting the female transition, increases the probability of a single male transition to marriage next period. Working part time in the past does not have any significant effect on males' transition from single to marriage. However, working part time or full time last period reduces the probability that a single female will transition to marriage next period, while working full time 2 year in the past reduces the probability that a single male transition to marriage next period. The age distribution of current children or the time spent with them do not have a significant effect on the transition probability of a single female, however, the older the second child of a single male the more likely he is to get marry next period.

The right hand panel of Table 9 shows that all the current choices of a single female increase the probability she will transition to married next period relative to choosing "no work-no birth-no time with children". For males all choices except those that involve a choice of not working while spending time with children (i.e. choices 4 and 7) increase the likelihood he will transition to marriage next period relative to not working while providing no parental time. In fact we find that if a single male chooses to work part time and supply low parental time he will transition to marriage next period with probability one.

Table 9 presents the logit coefficient estimates of the one period divorce rates. It shows that black females have a higher divorce rate than their white counterpart while there are no differences between the black and white males one period divorce rates. There is also no effect of a person's education on the one period divorce rate. For females the one period divorce rate is convex in age while age does not have any significant effect on the one period divorce rate of males. A similar patterns hold for the number of children. Table 9 also shows that if a female worked full time last period she is more likely to get divorce next period than a female who did not work or worked part time last period. Past work behavior does not have any significant effect on males' one period divorce rate. The age distribution of current children does not have any effect on female's one period divorce rate, however, the older a male's 4th child, the less likely he will get divorce next period. The time spent with current kids in the past or the number of female kids does not have any effect on the one period divorce rates of females. However, the more time a male spends with his 3rd child the higher the one period divorce rate while the more time he spends with his 4th child reduces the divorce rate. Overall it seems that if a male has four kids he is less likely to get divorced next period.

Table 9: Logit Coefficient Estimates Transition from Single to Married

Dependent Variable: Dummy equal one if married and zero otherwise

(Standard Error in parentheses)

State Variables			Choice Variables		
Variables	Female	Male	Choice	Female	Male
Black	-1.339	-1.952	2	1.365	0.951
	(0.066)	(0.168)		(0.132)	(0.289)
High School	0.300	0.172	3	1.005	1.774
	(0.101)	(0.153)		(0.092)	(0.134)
Some College	0.108	0.029	4	1.552	0.320
	(0.104)	(0.158)		(0.333)	(1.072)
College Graduate	-0.297	0.167	5	0.820	
	(0.109)	(0.157)		(0.205)	
Age	0.324	0.408	6	1.251	1.646
	(0.040)	(0.064)		(0.237)	(0.299)
Age Sq	-0.006	-0.007	7	1.249	0.622
	(0.001)	(0.001)		(0.162)	(1.063)
No. of Children	-0.338	1.849	8	1.303	1.410
	(0.205)	(0.412)		(0.240)	(1.115)
No. of Children Sq	0.078	-0.216	9	1.555	2.406
	(0.069)	(0.144)		(0.331)	(0.301)
Part time work (t-1)	-0.268	-0.128	10	1.183	
	(0.135)	(0.270)		(0.411)	
Part time work (t-2)	0.060	-0.399	11	1.210	
	(0.130)	(0.289)		(0.223)	
Part time work (t-3)	0.143	-0.201	12	1.754	
	(0.132)	(0.361)		(0.301)	
Part time work (t-4)	-0.105	-0.144	13	1.450	
	(0.136)	(0.358)		(0.209)	
Full time work (t-1)	-0.264	0.025	14	1.400	
	(0.102)	(0.159)		(0.243)	
Full time work (t-2)	0.166	-0.530	15	1.763	
	(0.106)	(0.178)		(0.431)	
Full time work (t-3)	-0.129	0.100	16	1.781	
	(0.113)	(0.207)		(0.309)	
Full time work (t-4)	-0.146	0.014			
	(0.101)	(0.189)			

Table 9 (cont'd): Logit Coefficient Estimates Transition from Single to Married

Variables	Female	Male			
Age of 1st Child	0.026	0.008			
	(0.018)	(0.032)			
Age of 2nd Child	0.007	-0.082			
	(0.029)	(0.050)			
Age of 3rd Child	0.030				
	(0.050)				
Age of 4th Child	0.170				
	(0.128)				
Time with 1st Child	-0.010	-0.013			
	(0.032)	(0.058)			
Time with 2nd Child	-0.020	-0.356			
	(0.044)	(0.116)			
Time with 3nd Child	-0.046				
	(0.070)				
Time with 4th Child	-0.316				
	(0.184)		Constant	-6.527	-9.457
No. of Female Children	-0.053	-0.111		(0.498)	(0.810)
	(0.073)	(0.179)	N	30,875	30,492

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997. Choice 5 for male is deterministic and is excluded; meaning if single male chooses to work part time and supply low child care hours he will get married next period with probability one.

Table 10: Logit Coefficient Estimates Transition from Married to Married

Dependent Variable: Dummy equal one if married and zero otherwise

(Standard Error in parentheses)

	State Variables					Choice Variables			
	Individual		Spouse			Individual		Spouse	
Variables	Female	Male	Female	Male	Choice	Female	Male	Female	Male
Black	-0.825	-0.397			2	-0.483	1.042	0.488	2.619
	(0.098)	(0.289)				(0.197)	(0.553)	(0.159)	(0.527)
High School	0.037	0.038	0.019	-0.407	3	-0.665	1.112	1.860	3.525
	(0.130)	(0.224)	(0.111)	(0.271)		(0.158)	(0.408)	(0.122)	(0.330)
Some College	-0.118	0.223	0.129	-0.610	4	-0.213	0.518	0.136	
	(0.137)	(0.240)	(0.121)	(0.284)		(0.514)	(1.085)	(0.248)	
College Graduate	0.161	0.431	0.576	-0.552	5	-0.034		0.012	3.508
	(0.164)	(0.258)	(0.146)	(0.313)		(0.224)		(0.253)	(0.345)
Age	-0.155	-0.047	0.190	-0.136	6	-0.041	0.673	2.114	3.875
	(0.067)	(0.140)	(0.053)	(0.169)		(0.238)	(0.434)	(0.163)	(0.456)
Age Square	0.003	0.000	-0.003	0.002	7	-0.461	-0.536	0.814	3.745
	(0.001)	(0.002)	(0.001)	(0.003)		(0.193)	(0.616)	(0.296)	(0.279)
No.of Children	-0.349	-0.637			8	-0.125	0.553	0.378	2.759
	(0.179)	(0.425)				(0.257)	(0.820)	(0.272)	(0.528)
No. of Children Sq	0.039	0.146			9	-0.269	0.894	1.654	3.020
	(0.053)	(0.150)				(0.285)	(0.451)	(0.164)	(0.769)
Part time work (t-1)	-0.207	0.480	0.037	1.024	10	-0.034			3.273
	(0.128)	(0.473)	(0.184)	(0.223)		(0.336)			(0.552)
Part time work (t-2)	0.121	-0.422	0.025	-0.496	11	0.463			2.273
	(0.136)	(0.403)	(0.202)	(0.219)		(0.232)			(0.220)
Part time work (t-3)	-0.126	0.295	0.277	-0.232	12	-0.063			2.728
	(0.144)	(0.429)	(0.234)	(0.208)		(0.248)			(0.320)
Part time work (t-4)	-0.140	-0.649	0.737	-0.283	13	-0.304			3.273
	(0.135)	(0.399)	(0.260)	(0.197)		(0.219)			(0.317)
Full time work (t-1)	-0.264	-0.098	-0.049	1.830	14	0.296			2.592
	(0.119)	(0.411)	(0.112)	(0.226)		(0.258)			(0.363)
Full time work (t-2)	0.163	-0.038	0.088	-1.028	15	-0.242			3.111
	(0.129)	(0.361)	(0.119)	(0.223)		(0.332)			(0.777)
Full time work (t-3)	-0.093	-0.045	0.213	-0.031	16	0.473			4.106
	(0.135)	(0.358)	(0.133)	(0.232)		(0.386)			(1.056)
Full time work (t-4)	0.138	-0.270	0.432	-0.490					
	(0.122)	(0.322)	(0.121)	(0.201)					

Table 10 (cont'd): Logit Coefficient Estimates Transition from Married to Married

	State Variables						
	Individual		Spouse				
Variables	Female	Male	Female	Male			
Age of 1st Child	-0.003	-0.021					
	(0.018)	(0.027)					
Age of 2nd Child	-0.003	-0.014					
	(0.025)	(0.031)					
Age of 3rd Child	-0.023	-0.096					
	(0.041)	(0.079)					
Age of 4th Child	0.076	0.226					
	(0.079)	(0.109)					
Time with 1st Child	-0.043	-0.033	0.088	-0.136			
	(0.031)	(0.041)	(0.029)	(0.048)			
Time with 2nd Child	0.052	0.072	-0.016	0.099			
	(0.038)	(0.063)	(0.036)	(0.053)			
Time with 3rd Child	0.010	-0.222	0.079	0.222			
	(0.062)	(0.109)	(0.060)	(0.129)			
Time with 4th Child	-0.054	0.771	0.045	-0.494			
	(0.092)	(0.378)	(0.171)	(0.144)	Constant	0.450	4.779
No. of Female Children	-0.046	-0.056				(0.819)	(1.811)
	(0.066)	(0.111)			N	23,694	14,740

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997. Individuals choice 5 and spouse choice 4 are deterministic for male and are excluded; meaning for a married male if these choices are chosen he will remain married next period with probability one.

Table 10 also shows that males' whose spouse has some college or a college degree are more likely to get divorced while the opposite is true for females. The older a female spouse the less likely she is to get divorce next period. A male whose spouse worked part or full time last period is less likely to get divorce next period relative to one with a spouse who did not work; the same is true for a female spouse who worked part or full time 4 years in the past. This pattern is reversed for males whose spouse worked full or part time 2 or 4 years in the past. Males whose spouse provide high parental time investment in the 1st and 4th child are more likely to get divorce next period.

Females who work part time, give birth, and do not provide any child care hours in the current period (i.e. Choice 4) are more likely to divorce next period. The same is true for females who work full time, do not give birth, and provide low child care hours in the current period (i.e. Choice 7) . The opposite is true for a female who does not work or give birth, but provides high child care hours (i.e. Choice 11). On the other hand, a male who works full or part time and provides no child care hours (i.e. Choices 2 and 3) has a lower probability of divorce next period relative to a male who does not work or provide any parental time investment. The same is true for a male who worked full time and provide high parental time investment (i.e. Choice 4). Again, we find that males that worked part time and provide low parental time never get divorce in our sample.

When it comes to the choices of females' spouse the patterns are not so clear. We find that a female whose spouse works full or part time and does not provide any child care (i.e. Choices 2 and 3) has a higher probability of remaining married next period relative to a female whose spouse does not work or provides parental time investment. The same is true for a female whose spouse works full time and provides some parental time investment (i.e. Choices 6 and 9) or does not work but provides high parental time investment (i.e. Choice 7). For males all spouse choices lead to a lower divorce rate relative to choosing no work, no birth, and no parental time.

3.4.1.4 Conditional Choice Probabilities of Single Females Table 11 presents the logit coefficient estimates of the conditional probability for single females. The excluded category is choice 1, which in not participating in the labor market, not giving birth, and not providing parental time investment. It shows that black females are less likely to choose choices 2, 3, 7, and 13; the first two involve working full or part time while not giving birth or investing time in children and the last two involve working full time while not giving birth and providing high or low parental time investment. On the other hand, black females are more likely to choose choices 4, 8, and 9; the predominant feature of these choices is giving birth. Therefore single black females are more likely to give birth than single white females.

It also shows that single female college graduates are less likely to choose choices 5, 8, 11, and 14 which involve not working. At the same time they are more likely to choose choices 3 and 7 which involve working full time, not giving birth, and providing no or low levels of parental time investment. While not as strong, a similar pattern holds for females with high school or some college education. The number of children increases the likelihood of any choice other than 1, at a decreasing rate. The same is true for all form of labor market experience.

Table 11 also shows that the older the 1st child of a single female, the more likely she chooses choice 1 relative to all the other choices, while the age of the 2nd child only has a significant positive effect on choice 3 (i.e. full time work and no birth or parental time investment) relative to the choice 1. The age of the 3rd child has a significant positive effect on choice 2 (i.e. part time work, no birth, and no parental time invest) and choice 4 (i.e. full time work, birth, and no parental time investment); however, the effect on choice 4 is much greater than on choice 2. The age of the 4th child has a significant positive effect on choices 2 and 4, which is similar to the effect of the age of the 3th child. Unlike the effect of the age of the 3rd child, the effect of the age of the 4th child on the likelihood of choices 8, 9, 10, 14, 15, and 16 is negative. The predominant features of all these choice are giving birth and providing positive parental investment. Past time investment in the 1st child has a positive effect on the likelihood of choice 5 through 16 relative to choice 1; these are all choices that involve providing positive amount of parental time investment. The only negative effect of past parental time investment in the 1st child is on choice 4, which is full time work, giving birth, and providing no parental time investment. Past parental investment in the 2nd child has a significant negative effect on the likelihood of choices 3, 5, and 6 relative to choice 1, all involving not giving birth. The effects of parental time investment in the 3th child are similar to those of parental time investment in the 1st child except they are not as significant. There are no clear patterns to the effect of parental time investment in the 4th child (there are both negative and positive effects on different aspects of the choices). Finally, the number of female children reduces the likelihood of choice 9 and 12, which all involve working part time with positive parental time investment.

Table 11: Logit Coefficient of Conditional Choice Probability for Single Female

(Standard Error in parenthesis; Choice 1 is the excluded class)

	Choice														
Variables	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Black	-0.221	-0.508	1.101	0.139	-0.063	-0.444	0.497	0.493	0.307	-0.222	-0.240	-0.272	0.001	0.287	0.271
	(0.092)	(0.061)	(0.250)	(0.105)	(0.127)	(0.086)	(0.150)	(0.263)	(0.230)	(0.126)	(0.176)	(0.117)	(0.168)	(0.335)	(0.239)
High Sch.	-0.236	0.742	1.291	-0.182	0.245	0.628	-0.717	0.420	0.289	-0.182	0.239	0.534	-0.391	-0.118	0.798
	(0.171)	(0.135)	(0.615)	(0.149)	(0.225)	(0.175)	(0.176)	(0.458)	(0.434)	(0.170)	(0.297)	(0.221)	(0.219)	(0.501)	(0.440)
Some Col.	-0.069	0.800	1.098	-0.193	0.534	0.808	-1.386	0.220	0.445	-0.591	0.436	0.513	-0.892	-0.372	0.323
	(0.169)	(0.135)	(0.624)	(0.162)	(0.229)	(0.179)	(0.212)	(0.469)	(0.441)	(0.201)	(0.324)	(0.232)	(0.255)	(0.545)	(0.462)
College	-0.042	0.828	0.395	-0.946	-0.000	0.639	-2.104	-0.333	0.536	-1.075	0.584	0.188	-2.076	-0.290	-0.444
	(0.174)	(0.137)	(0.680)	(0.245)	(0.287)	(0.197)	(0.320)	(0.552)	(0.498)	(0.311)	(0.359)	(0.269)	(0.410)	(0.600)	(0.558)
Age	0.390	0.517	1.002	0.209	0.460	0.186	0.949	0.861	0.577	0.026	-0.024	0.097	0.608	0.727	0.699
	(0.039)	(0.022)	(0.178)	(0.066)	(0.082)	(0.045)	(0.151)	(0.312)	(0.207)	(0.075)	(0.106)	(0.068)	(0.142)	(0.307)	(0.217)
Age Sq	-0.006	-0.008	-0.018	-0.003	-0.007	-0.003	-0.018	-0.017	-0.011	-0.000	0.000	-0.002	-0.010	-0.014	-0.012
	(0.001)	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.003)	(0.006)	(0.004)	(0.001)	(0.002)	(0.001)	(0.003)	(0.006)	(0.004)
No.of kids	1.474	1.222	3.740	8.462	8.800	8.238	3.662	2.272	9.317	8.270	8.976	8.207	3.523	1.798	2.715
	(0.442)	(0.367)	(0.910)	(0.376)	(0.492)	(0.372)	(0.456)	(0.844)	(0.767)	(0.420)	(0.584)	(0.453)	(0.432)	(0.949)	(0.630)
No. of kids Sq	-0.328	-0.387	-1.072	-2.122	-2.235	-2.123	-0.928	-0.101	-2.522	-2.057	-2.321	-2.175	-0.845	-0.802	-0.818
	(0.192)	(0.172)	(0.592)	(0.155)	(0.177)	(0.151)	(0.190)	(0.406)	(0.292)	(0.158)	(0.192)	(0.168)	(0.169)	(0.389)	(0.232)
Part work (t-1)	3.812	3.183	2.165	1.351	2.764	2.933	1.081	1.977	2.885	1.344	2.508	2.733	1.555	2.361	2.891
	(0.179)	(0.142)	(0.482)	(0.219)	(0.238)	(0.219)	(0.329)	(0.468)	(0.464)	(0.250)	(0.289)	(0.265)	(0.350)	(0.529)	(0.375)
Part work (t-2)	1.579	0.934	0.477	0.715	1.328	0.860	1.833	1.443	1.210	0.375	0.989	0.784	0.224	0.596	0.727
	(0.267)	(0.223)	(0.629)	(0.259)	(0.275)	(0.251)	(0.349)	(0.601)	(0.518)	(0.292)	(0.329)	(0.296)	(0.462)	(0.698)	(0.523)
Part work (t-3)	0.692	0.447	0.959	0.266	0.622	0.239	-0.028	-0.851	0.180	0.151	0.433	0.301	-0.161	0.470	0.490
	(0.299)	(0.251)	(0.517)	(0.270)	(0.293)	(0.266)	(0.414)	(1.010)	(0.504)	(0.295)	(0.339)	(0.304)	(0.525)	(0.669)	(0.565)
Part work (t-4)	0.519	0.092	-0.621	-0.141	0.415	0.394	0.790	-1.426	0.326	-0.018	0.740	0.110	-0.092	1.374	0.103
	(0.320)	(0.278)	(0.783)	(0.300)	(0.312)	(0.285)	(0.421)	(1.037)	(0.479)	(0.323)	(0.355)	(0.321)	(0.479)	(0.552)	(0.543)
Full work (t-1)	3.950	5.018	3.334	1.567	3.399	5.007	1.214	2.705	4.195	1.283	2.675	4.327	1.214	3.039	3.561
	(0.169)	(0.104)	(0.410)	(0.208)	(0.221)	(0.186)	(0.326)	(0.412)	(0.421)	(0.256)	(0.293)	(0.227)	(0.394)	(0.551)	(0.303)
Full work (t-2)	0.590	0.788	0.764	0.020	0.228	0.773	1.261	1.253	1.298	-0.304	0.015	0.799	0.420	-0.528	0.726
	(0.220)	(0.160)	(0.455)	(0.210)	(0.237)	(0.191)	(0.320)	(0.528)	(0.442)	(0.257)	(0.324)	(0.233)	(0.367)	(0.566)	(0.428)
Full work (t-3)	0.568	0.605	0.968	0.048	0.286	0.536	-0.397	0.947	0.823	-0.030	0.316	0.296	0.370	0.551	0.726
	(0.273)	(0.216)	(0.461)	(0.253)	(0.281)	(0.234)	(0.395)	(0.543)	(0.407)	(0.287)	(0.334)	(0.271)	(0.414)	(0.520)	(0.480)
Full work (t-4)	0.241	0.324	0.308	0.002	0.263	0.480	0.348	-0.453	0.327	0.114	0.514	0.366	-0.078	1.255	-0.422
	(0.257)	(0.212)	(0.398)	(0.237)	(0.260)	(0.223)	(0.419)	(0.515)	(0.376)	(0.264)	(0.318)	(0.253)	(0.422)	(0.567)	(0.438)

Table 11 (cont'd): Logit Coefficient of Conditional Choice Probability for Single Female

(Standard Error in parenthesis; Choice 1 is the excluded class)

	Choice														
Variables	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Age of 1st kid	-0.101	-0.087	-0.078	-0.136	-0.191	-0.121	-0.203	0.004	-0.255	-0.246	-0.199	-0.198	-0.477	-0.279	-0.493
	(0.033)	(0.022)	(0.048)	(0.026)	(0.029)	(0.022)	(0.060)	(0.100)	(0.060)	(0.035)	(0.046)	(0.030)	(0.107)	(0.129)	(0.152)
Age of 2nd kid	0.078	0.116	0.007	-0.003	0.035	-0.008	0.027	-0.141	0.059	-0.033	-0.073	-0.034	-0.059	-0.066	0.245
	(0.051)	(0.039)	(0.096)	(0.037)	(0.042)	(0.038)	(0.095)	(0.184)	(0.081)	(0.043)	(0.059)	(0.044)	(0.190)	(0.181)	(0.226)
Age of 3rd kid	0.112	0.044	-9.412	0.030	0.020	0.058	-0.081	-1.580	0.170	0.056	0.051	-0.068	-0.491	0.022	0.006
	(0.071)	(0.065)	(2.788)	(0.060)	(0.069)	(0.064)	(0.216)	(1.016)	(0.135)	(0.071)	(0.087)	(0.078)	(0.890)	(0.379)	(0.388)
Age of 4th kid	1.442	-0.686	5.923	-0.188	-0.167	-0.087	-2.061	-5.752	-3.577	-0.293	-0.058	-0.047	-3.450	-4.869	-6.196
	(0.384)	(0.367)	(1.672)	(0.392)	(0.396)	(0.388)	(0.749)	(2.923)	(1.663)	(0.395)	(0.408)	(0.402)	(1.634)	(1.686)	(2.968)
Time 1st kid	-0.066	0.036	-0.247	0.434	0.450	0.414	0.388	0.220	0.399	0.766	0.735	0.710	0.865	0.911	0.745
	(0.100)	(0.069)	(0.148)	(0.064)	(0.069)	(0.065)	(0.089)	(0.131)	(0.095)	(0.071)	(0.084)	(0.072)	(0.112)	(0.196)	(0.150)
Time 2nd kid	0.017	-0.109	-0.126	-0.178	-0.214	-0.110	-0.015	-0.028	-0.136	-0.038	0.005	-0.010	0.028	0.174	-0.100
	(0.128)	(0.098)	(0.389)	(0.087)	(0.095)	(0.090)	(0.131)	(0.271)	(0.136)	(0.093)	(0.109)	(0.097)	(0.181)	(0.215)	(0.299)
Time 3nd kid	-0.461	-0.153	-0.708	0.440	0.429	0.371	0.356	-0.492	0.486	0.431	0.672	0.716	0.621	0.635	0.372
	(0.272)	(0.157)	(0.797)	(0.132)	(0.149)	(0.138)	(0.208)	(0.564)	(0.227)	(0.147)	(0.161)	(0.151)	(0.642)	(0.307)	(0.411)
Time 4th kid	-7.249	2.360	4.533	3.624	3.838	3.582	-4.556	-3.899	0.868	3.809	3.674	3.679	-4.155	-3.013	-5.094
	(2.255)	(1.388)	(1.999)	(1.388)	(1.389)	(1.385)	(1.399)	(2.457)	(1.499)	(1.387)	(1.397)	(1.389)	(1.434)	(1.285)	(2.068)
Female kids	-0.025	0.010	-0.272	0.056	-0.064	-0.099	-0.295	-1.270	-0.143	-0.036	-0.358	-0.150	-0.218	-0.233	0.214
	(0.211)	(0.152)	(0.359)	(0.140)	(0.157)	(0.144)	(0.202)	(0.420)	(0.225)	(0.147)	(0.180)	(0.154)	(0.208)	(0.356)	(0.300)
Constant	-9.686	-11.047	-21.147	-9.696	-16.006	-11.057	-15.932	-17.205	-18.406	-7.992	-9.662	-10.838	-13.005	-15.908	-16.075
	(0.558)	(0.323)	(2.320)	(0.926)	(1.280)	(0.709)	(1.750)	(3.721)	(3.005)	(1.054)	(1.618)	(1.024)	(1.746)	(3.630)	(2.734)
N	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997.

3.4.1.5 Conditional Choice Probabilities of Single Males

Table 12 presents the logit coefficient estimates of the conditional choice probability for single males. It shows that black males are less likely than white males to choose choices 3, 4, 5, and 9 relative to the choice 1 (i.e. not working and providing parental time investment). It seems black males are less likely to specialized in parental time investment than white males and they are less likely to work full time.

Table 12 also shows that a college educated and high school graduate single males are more likely overall to work full time than single men with only some college. College graduates are more likely to make choices 3 and 5; these choices involve either full time work with no parental time investment or part time work with low parental time investment. A similar pattern holds for high school graduates or some college. On the other hand college graduate is less likely than single male with less than a high school education to choose choices 4, 7, and 8; these choices involve specialization in parental time investment to some extent. Similar patterns hold for high school graduate and some college. Similar to single females, the number of children increases the likelihood of single males making choices 4 through 9 relative to choice 1. All these choices involve providing some parental time investment. Therefore even single males with child are more likely to invest time in their children. The only negative effects of any type of labor market experience are on choices 4, 5, and 7; these are all choices that involve not working or working part time with low parental time investment. Therefore as with single female's labor market experience increases the likelihood of continue labor market participation. The only positive effect of the age distribution of kids on the choices of single males is the positive effect of the age of the 1st child on the probability of full time work while providing low parental time investment. Finally the number of female children increases the likelihood that a single male would choose choices 3, 5, 6, and 9; that is either working full time while not providing any parental time investment or working and working with some parental time investment.

Table 12: Logit Coefficient of Conditional Choice Probability for Single Male

(Standard Error in parenthesis; Choice 1 is the excluded class)

	Choice							
Variables	2	3	4	5	6	7	8	9
Black	0.162	-0.392	-11.687	-1.034	-0.627	1.080	0.020	-1.085
	(0.096)	(0.061)	(1.467)	(0.803)	(0.408)	(0.783)	(0.908)	(0.399)
High Sch.	-0.304	0.257	-0.352	10.887	0.490	0.664	2.131	0.792
	(0.143)	(0.091)	(1.050)	(1.535)	(0.376)	(0.924)	(1.544)	(0.383)
Some Col.	-0.207	0.199	-1.564	9.350	0.050	0.257	1.003	-0.119
	(0.150)	(0.095)	(1.424)	(1.896)	(0.384)	(1.377)	(1.613)	(0.401)
College	-0.176	0.416	-10.694	9.523	0.638	-9.201	-9.494	0.522
	(0.158)	(0.096)	(1.930)	(1.560)	(0.401)	(2.613)	(1.843)	(0.405)
Age	0.747	0.878	4.777	0.598	1.231	0.175	2.905	1.200
	(0.070)	(0.038)	(2.284)	(0.419)	(0.194)	(0.423)	(1.173)	(0.170)
Age Sq	-0.013	-0.015	-0.066	-0.010	-0.020	-0.003	-0.040	-0.018
	(0.001)	(0.001)	(0.032)	(0.006)	(0.003)	(0.007)	(0.017)	(0.002)
No.of kids	-0.344	-0.639	9.951	8.270	5.007	13.350	18.071	5.533
	(1.133)	(0.988)	(2.536)	(2.767)	(1.070)	(1.945)	(1.833)	(1.420)
No. of kids Sq	-0.095	-0.053	-1.986	-1.884	-1.255	-3.021	-6.542	-1.431
	(0.205)	(0.170)	(0.834)	(0.883)	(0.249)	(0.718)	(0.830)	(0.434)
Part work (t-1)	4.217	3.321	3.387	12.425	3.291	-12.458	3.564	4.093
	(0.198)	(0.154)	(1.507)	(1.619)	(1.491)	(2.129)	(1.330)	(0.842)
Part work (t-2)	1.625	0.864	1.918	-8.906	19.273	2.239	-1.285	1.877
	(0.340)	(0.306)	(1.264)	(1.505)	(3.549)	(1.254)	(1.860)	(1.079)
Part work (t-3)	-0.070	-0.731	2.332	-2.607	-1.128	0.017	0.551	-0.854
	(0.405)	(0.359)	(1.106)	(1.170)	(0.901)	(1.716)	(1.581)	(0.929)
Part work (t-4)	0.788	0.318	-1.086	12.434	1.755	1.296	2.003	1.473
	(0.446)	(0.382)	(1.439)	(1.280)	(0.911)	(1.810)	(1.483)	(0.783)
Full work (t-1)	4.397	5.075	-0.887	10.881	5.668	0.274	3.195	4.791
	(0.169)	(0.101)	(1.559)	(1.238)	(1.255)	(0.994)	(1.357)	(0.735)
Full work (t-2)	0.787	1.079	2.434	1.101	19.181	-0.119	0.431	2.194
	(0.255)	(0.203)	(1.739)	(1.012)	(3.549)	(1.891)	(1.558)	(0.874)
Full work (t-3)	0.205	0.443	-0.200	-2.632	-0.460	-0.928	0.525	-0.056
	(0.350)	(0.284)	(1.413)	(1.324)	(0.800)	(1.636)	(1.624)	(0.811)
Full work (t-4)	0.741	0.599	-2.839	8.379	1.543	-1.522	0.705	1.187
	(0.338)	(0.283)	(0.981)	(1.460)	(0.754)	(1.048)	(1.258)	(0.650)
Age of 1st kid	0.100	0.188	0.064	0.006	0.320	-0.042	0.136	0.162
	(0.158)	(0.135)	(0.267)	(0.200)	(0.138)	(0.185)	(0.146)	(0.139)
Age of 2nd kid	0.050	-0.063	-0.504	-0.029	-0.205	-0.302	0.175	-0.168
	(0.133)	(0.123)	(0.352)	(0.170)	(0.128)	(0.341)	(0.187)	(0.129)
Female kids	1.402	1.793	-0.329	1.404	1.247	1.091	-1.029	1.446
	(0.831)	(0.672)	(1.864)	(0.717)	(0.667)	(0.859)	(1.291)	(0.658)
Constant	-14.516	-14.955	-94.644	-44.118	-46.713	-13.193	-68.683	-29.242
	(0.910)	(0.481)	(40.775)	(6.828)	(0.000)	(6.150)	(21.834)	(2.969)
N	35,939	35,939	35,939	35,939	35,939	35,939	35,939	35,939

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997.

3.4.1.6 Best Response Functions Unlike single individuals, married couples are engaged in a non-corporative game of complete information, therefore we have to estimate the best response function of each spouse. These best response functions do not only depend on the individual's state space but also on the state space and choices of their spouses.

Females' Ex-ante Best Response Probabilities

Table 13 presents the logit coefficient estimates of ex-ante conditional best response probabilities of a married female. It shows that the behavior of single black females and married black females differs significantly. Specifically, married black females are less likely to choose 3, 5, 6, 7, 11, 12, 13 and 14 relative to their white counterparts. The first choice is working full time while doing nothing else; the next three choices (i.e. choices 5, 6, and 7) involve not giving birth while providing low parental time investment; and the last four (i.e. choices 11, 12, 13, and 14) involve high time investment while either giving birth, working, or doing nothing. So while they behave differently from white married females it is hard to make any generalization as the choices include different combinations of work, birth, and parental time investment, however, overall black married females are less likely to make choices involving high parental time investment relative to white married women. Similar to single female, college educated married females are more likely to choose almost all other choices relatives to choice 1. This pattern is similar for high school graduates and some college education. The same is true for the effect of the number of children. Again all types of labor market experiences make it more likely to work in the current period.

Table 13 also shows that the age of the 1st child has a significant negative effect on the likelihood of choices 8, 11, 14, 15, and 16; most of these choices involve giving birth in the current period. The effects of the age distribution of older children are not as striking as those of the age of the 1st child. Parental time investment in the 1st child has a significant positive effect on the likelihood of choices 5 through 16; therefore past parental time investment in the 1st child leads to higher likelihood of current parental time investment. The pattern is reversed for parental time investment in the 2nd child, in fact the likelihood of the choices relative to doing nothing, except choice 2 which is statistically insignificant, increases in the time invested in the second child. This may be because most families have only 2 children. This pattern is repeated for parental time investment in the 3rd and 4th child.

The second panel of Table 13 presents the effect of spouse's characteristics on the ex-ante conditional best response of married female. If a female's spouse is a college graduate, the female has a higher likelihood of choosing 3, 5, 6, 8, 11, and 14. As usual similar patterns hold for high school or some college education. Therefore education of the spouse increases the likelihood of specialization either in the labor market or at home. Spouse's labor market experience has the opposite effect on the likelihood choices relative to the female's own labor market. All else equal, the more labor market experience a female's spouse has, the more likely that the female will choose not to work. The more parental investment a female's spouse made in their 1st child, the lower the likelihood of the female choosing 11 through 16. These are all choices involving high parental time investment. This shows that fathers' parental investment seems to be a substitute for mothers' parental investment. A similar pattern holds for the spouse's parental time investment in the 3rd child, except that there is also a reduced likelihood of the female choosing choices 5,6, and 7. The additional choices involve low parental time investment of the female. The effect of the 4th child is similar to those above except that higher spouse parental time investment in the 4th child increases the likelihood of female choosing not to work while giving birth and providing high parental time investment.

The final panel of Table 13 presents the reaction function of spouse's choices on the female ex-ante probability of choices. It shows that if the spouse choose to work part time (i.e. spouse choices 2, 5, and 8) the female is more likely to work. If the spouse works full time (i.e. spouse choices 3, 6, and 9) the female is still more likely to work but is also more likely to give birth or provide positive parental time investment. If the spouse chooses not to work and provide low parental time investment, the female is less likely to choose 2, 4, and 11. These choices involve either not providing parental time investment and work full time (whether the female chooses to give birth or not) or provide high time investment in children and not work.

Table 13: Logit Coefficient of Best Response Probability for Married Female

(Standard Error in parenthesis; Choice 1 is the excluded class)

Individual Variables	Choice 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Black	-0.189	-0.407	0.245	-0.549	-0.679	-0.362	-0.375	0.001	-0.211	-1.284	-0.893	-0.674	-0.762	-0.395	-0.228
	(0.156)	(0.122)	(0.303)	(0.181)	(0.196)	(0.156)	(0.207)	(0.223)	(0.226)	(0.193)	(0.223)	(0.176)	(0.200)	(0.278)	(0.226)
High Sch.	0.727	0.926	0.111	0.370	0.524	0.877	0.309	0.625	0.565	0.372	0.669	0.578	0.414	1.064	0.377
	(0.198)	(0.147)	(0.551)	(0.158)	(0.190)	(0.160)	(0.192)	(0.327)	(0.298)	(0.150)	(0.183)	(0.167)	(0.169)	(0.352)	(0.246)
Some Col.	1.009	1.282	0.604	0.341	0.595	0.900	0.134	0.958	0.800	0.195	0.733	0.575	0.384	1.156	0.632
	(0.208)	(0.159)	(0.547)	(0.178)	(0.205)	(0.174)	(0.222)	(0.337)	(0.310)	(0.171)	(0.203)	(0.184)	(0.188)	(0.376)	(0.265)
College	1.382	1.547	0.960	0.226	0.681	1.014	0.380	1.075	1.111	-0.078	0.832	0.674	0.419	1.244	0.801
	(0.231)	(0.179)	(0.556)	(0.209)	(0.230)	(0.198)	(0.251)	(0.355)	(0.337)	(0.199)	(0.233)	(0.210)	(0.217)	(0.403)	(0.295)
Age	-0.170	-0.022	0.598	-0.106	-0.169	-0.296	0.371	0.347	0.253	-0.107	-0.347	-0.343	0.386	-0.092	-0.254
	(0.075)	(0.056)	(0.235)	(0.077)	(0.087)	(0.067)	(0.114)	(0.146)	(0.144)	(0.077)	(0.093)	(0.075)	(0.105)	(0.154)	(0.129)
Age Sq	0.003	0.001	-0.009	0.002	0.003	0.005	-0.007	-0.007	-0.005	0.002	0.006	0.006	-0.008	0.001	0.004
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
No.of kids	0.288	0.692	1.298	3.505	3.275	3.380	1.098	0.615	3.144	3.067	2.961	2.909	0.306	0.323	-0.077
	(0.284)	(0.201)	(0.368)	(0.208)	(0.209)	(0.201)	(0.247)	(0.289)	(0.224)	(0.206)	(0.214)	(0.206)	(0.232)	(0.282)	(0.262)
Part work (t-1)	1.416	0.831	1.058	-0.105	1.884	2.142	0.328	1.504	1.628	-0.493	1.351	1.442	-0.148	1.304	1.035
	(0.149)	(0.129)	(0.462)	(0.145)	(0.149)	(0.150)	(0.164)	(0.206)	(0.265)	(0.141)	(0.147)	(0.157)	(0.149)	(0.196)	(0.207)
Part work (t-2)	0.494	0.100	0.052	0.228	0.516	0.380	-0.166	0.524	0.354	0.051	0.606	0.354	0.285	0.723	0.528
	(0.186)	(0.157)	(0.491)	(0.170)	(0.174)	(0.166)	(0.221)	(0.258)	(0.292)	(0.165)	(0.175)	(0.177)	(0.179)	(0.239)	(0.258)
Part work (t-3)	0.311	-0.130	-0.418	-0.102	0.366	0.139	0.042	0.386	0.439	-0.122	0.359	0.107	0.004	0.317	0.114
	(0.218)	(0.187)	(0.497)	(0.198)	(0.200)	(0.192)	(0.250)	(0.270)	(0.292)	(0.193)	(0.202)	(0.201)	(0.209)	(0.271)	(0.276)
Part work (t-4)	-0.090	-0.199	0.147	-0.322	0.118	-0.144	-0.048	0.144	0.019	-0.201	-0.085	0.027	0.209	0.273	-0.088
	(0.233)	(0.183)	(0.469)	(0.199)	(0.200)	(0.190)	(0.259)	(0.269)	(0.271)	(0.193)	(0.202)	(0.200)	(0.209)	(0.280)	(0.284)
Full work (t-1)	1.710	2.860	2.799	-0.234	1.673	3.764	-0.250	1.496	2.655	-1.135	0.713	2.703	-0.772	0.704	1.695
	(0.148)	(0.119)	(0.363)	(0.157)	(0.161)	(0.149)	(0.187)	(0.220)	(0.244)	(0.163)	(0.169)	(0.154)	(0.171)	(0.216)	(0.196)
Full work (t-2)	0.300	0.656	0.768	0.054	-0.027	0.628	0.386	0.813	1.165	-0.042	-0.066	0.533	0.352	0.698	1.197
	(0.189)	(0.151)	(0.420)	(0.181)	(0.185)	(0.165)	(0.208)	(0.259)	(0.253)	(0.176)	(0.194)	(0.176)	(0.181)	(0.252)	(0.225)
Full work (t-3)	0.041	0.121	-0.383	0.100	0.338	0.517	0.098	0.319	0.895	0.040	0.338	0.501	0.007	-0.004	0.281
	(0.216)	(0.176)	(0.427)	(0.203)	(0.205)	(0.188)	(0.247)	(0.277)	(0.267)	(0.195)	(0.210)	(0.197)	(0.210)	(0.277)	(0.254)
Full work (t-4)	0.022	0.098	0.669	-0.110	0.093	0.332	0.220	0.176	0.305	0.106	0.075	0.551	0.222	0.694	0.637
	(0.198)	(0.153)	(0.358)	(0.178)	(0.180)	(0.165)	(0.227)	(0.252)	(0.226)	(0.171)	(0.185)	(0.173)	(0.186)	(0.245)	(0.225)
Age of 1st kid	0.037	0.073	0.095	-0.061	-0.075	0.007	-0.116	-0.053	-0.046	-0.112	-0.068	-0.062	-0.333	-0.287	-0.258
	(0.047)	(0.038)	(0.051)	(0.039)	(0.040)	(0.038)	(0.049)	(0.056)	(0.044)	(0.039)	(0.041)	(0.039)	(0.057)	(0.076)	(0.053)
Age of 2nd kid	0.026	0.022	-0.053	-0.101	-0.058	-0.086	0.000	-0.136	-0.055	-0.146	-0.141	-0.136	-0.015	-0.094	-0.007
	(0.056)	(0.043)	(0.108)	(0.043)	(0.044)	(0.042)	(0.070)	(0.090)	(0.060)	(0.045)	(0.047)	(0.043)	(0.079)	(0.110)	(0.080)
Age of 3rd kid	-0.032	-0.008	-40.662	-0.112	-0.132	-0.143	-0.176	-0.067	-0.095	-0.255	-0.234	-0.190	-0.324	-1.219	-0.423
	(0.111)	(0.080)	(0.000)	(0.078)	(0.075)	(0.072)	(0.125)	(0.204)	(0.130)	(0.084)	(0.085)	(0.075)	(0.166)	(0.531)	(0.254)
Age of 4th kid	-0.281	-0.221	-3.786	-0.270	-0.234	-0.257	-17.703	-13.605	-22.840	-0.077	-0.186	-0.176	-0.724	-3.793	-7.562
	(0.188)	(0.211)	(0.000)	(0.154)	(0.155)	(0.153)	(0.967)	(1.363)	(0.899)	(0.160)	(0.166)	(0.162)	(0.568)	(1.337)	(0.974)
Time 1st kid	-0.071	-0.142	-0.133	0.590	0.616	0.619	0.582	0.565	0.680	0.906	0.878	0.896	1.034	1.057	1.073
	(0.128)	(0.085)	(0.170)	(0.081)	(0.082)	(0.081)	(0.091)	(0.107)	(0.088)	(0.082)	(0.083)	(0.082)	(0.090)	(0.104)	(0.093)
Time 2nd kid	-0.095	-0.265	-0.397	-0.431	-0.418	-0.424	-0.303	-0.024	-0.654	-0.278	-0.200	-0.202	-0.035	0.049	-0.010
	(0.099)	(0.071)	(0.192)	(0.068)	(0.069)	(0.067)	(0.104)	(0.114)	(0.096)	(0.068)	(0.071)	(0.068)	(0.086)	(0.115)	(0.102)

Table 13 (cont'd): Logit Coefficient of Best Response Probability for Married Female

(Standard Error in parenthesis; Choice 1 is the excluded class)

	Choice														
Individual Variables	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Time 3nd kid	0.011	-0.256	-29.186	-0.256	-0.250	-0.232	0.138	-0.033	-0.205	-0.025	-0.062	-0.037	0.507	0.977	0.474
	(0.218)	(0.178)	(0.000)	(0.153)	(0.152)	(0.149)	(0.221)	(0.203)	(0.198)	(0.154)	(0.155)	(0.152)	(0.189)	(0.327)	(0.216)
Time 4th kid	0.650	0.659	-2.247	0.045	0.050	-0.094	-13.526	-13.969	-21.740	0.067	0.113	0.069	-0.048	-8.441	-7.429
	(0.242)	(0.311)	(0.000)	(0.201)	(0.207)	(0.207)	(1.026)	(1.406)	(0.000)	(0.205)	(0.213)	(0.215)	(0.470)	(1.232)	(1.143)
Female kids	0.185	-0.162	0.029	-0.072	0.010	-0.082	0.065	-0.334	-0.092	-0.100	-0.143	-0.131	-0.058	-0.082	0.058
	(0.211)	(0.160)	(0.321)	(0.156)	(0.157)	(0.153)	(0.191)	(0.222)	(0.175)	(0.154)	(0.157)	(0.155)	(0.171)	(0.213)	(0.202)
Spouse Variables															
High Sch.	0.138	0.381	0.718	0.404	0.457	0.410	0.118	0.274	0.449	0.369	0.372	0.331	0.152	-0.124	0.186
	(0.161)	(0.120)	(0.537)	(0.154)	(0.178)	(0.140)	(0.182)	(0.272)	(0.248)	(0.146)	(0.171)	(0.151)	(0.160)	(0.258)	(0.217)
Some Col.	0.411	0.448	0.999	0.333	0.472	0.380	0.142	0.503	0.476	0.375	0.223	0.135	0.214	0.064	0.140
	(0.173)	(0.133)	(0.547)	(0.173)	(0.192)	(0.154)	(0.205)	(0.288)	(0.260)	(0.163)	(0.190)	(0.167)	(0.177)	(0.277)	(0.237)
College	0.313	0.334	0.653	0.483	0.550	0.226	0.491	0.564	0.367	0.447	0.314	-0.271	0.373	0.338	-0.339
	(0.187)	(0.145)	(0.581)	(0.187)	(0.205)	(0.168)	(0.227)	(0.299)	(0.282)	(0.178)	(0.205)	(0.184)	(0.196)	(0.284)	(0.267)
Age	0.099	0.066	0.224	-0.062	0.007	0.004	-0.019	-0.048	-0.033	-0.012	0.099	-0.105	0.003	0.068	0.109
	(0.065)	(0.047)	(0.170)	(0.066)	(0.079)	(0.058)	(0.092)	(0.104)	(0.121)	(0.066)	(0.080)	(0.066)	(0.080)	(0.133)	(0.103)
Age Sq	-0.002	-0.001	-0.005	0.001	-0.001	-0.000	-0.000	-0.000	-0.000	-0.000	-0.002	0.001	-0.001	-0.002	-0.003
	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Part work (t-1)	-0.609	-0.810	-0.171	0.064	-0.580	-0.438	0.014	-1.003	0.027	-0.125	-0.515	-0.166	-0.194	-0.655	-0.417
	(0.272)	(0.213)	(0.518)	(0.288)	(0.322)	(0.262)	(0.339)	(0.464)	(0.422)	(0.287)	(0.377)	(0.298)	(0.315)	(0.503)	(0.459)
Part work (t-2)	-0.696	-0.514	-0.534	-0.365	-0.347	-0.323	-0.195	-0.189	-0.764	-0.325	-0.707	-0.281	-0.897	0.031	-1.755
	(0.316)	(0.237)	(0.540)	(0.295)	(0.327)	(0.271)	(0.370)	(0.390)	(0.433)	(0.297)	(0.372)	(0.303)	(0.377)	(0.453)	(0.602)
Part work (t-3)	-0.075	-0.426	-0.973	-0.487	-0.794	-0.437	-0.672	-0.423	-0.238	-0.607	-0.846	-0.632	-0.552	-0.663	0.211
	(0.331)	(0.256)	(0.649)	(0.350)	(0.359)	(0.304)	(0.409)	(0.446)	(0.436)	(0.328)	(0.387)	(0.330)	(0.385)	(0.531)	(0.449)
Part work (t-4)	-0.142	-0.067	-0.228	-0.311	-0.046	-0.251	-0.375	-0.098	-0.647	-0.713	-0.426	-0.660	-0.009	-0.477	-0.761
	(0.356)	(0.275)	(0.639)	(0.337)	(0.353)	(0.304)	(0.445)	(0.476)	(0.469)	(0.338)	(0.378)	(0.334)	(0.348)	(0.541)	(0.475)
Full work (t-1)	-0.676	-1.057	-0.609	0.020	-0.327	-0.334	0.214	-0.251	0.144	0.207	0.093	0.010	0.237	-0.078	0.071
	(0.143)	(0.116)	(0.345)	(0.176)	(0.194)	(0.150)	(0.187)	(0.241)	(0.265)	(0.172)	(0.228)	(0.179)	(0.166)	(0.267)	(0.225)
Full work (t-2)	-0.181	-0.205	-0.364	-0.032	0.200	0.110	0.155	0.067	-0.090	0.327	0.215	0.305	0.100	0.145	-0.372
	(0.168)	(0.131)	(0.337)	(0.179)	(0.195)	(0.155)	(0.192)	(0.242)	(0.233)	(0.174)	(0.225)	(0.181)	(0.179)	(0.255)	(0.220)
Full work (t-3)	-0.075	-0.136	-0.378	0.197	-0.334	0.010	-0.271	-0.471	0.066	-0.074	-0.191	-0.191	0.211	-0.196	0.205
	(0.205)	(0.157)	(0.417)	(0.202)	(0.213)	(0.177)	(0.228)	(0.274)	(0.260)	(0.188)	(0.227)	(0.193)	(0.204)	(0.277)	(0.251)
Full work (t-4)	-0.175	-0.123	-0.293	-0.248	-0.058	-0.308	-0.269	-0.002	-0.339	-0.405	-0.360	-0.497	-0.481	-0.521	-0.731
	(0.191)	(0.144)	(0.381)	(0.181)	(0.194)	(0.161)	(0.220)	(0.261)	(0.234)	(0.171)	(0.200)	(0.173)	(0.187)	(0.263)	(0.224)
Time 1st kid	-0.004	0.021	0.105	-0.063	-0.113	-0.106	-0.108	-0.080	-0.069	-0.204	-0.178	-0.173	-0.201	-0.358	-0.230
	(0.099)	(0.078)	(0.129)	(0.077)	(0.077)	(0.075)	(0.089)	(0.093)	(0.081)	(0.077)	(0.078)	(0.076)	(0.081)	(0.094)	(0.086)
Time 2nd kid	-0.052	0.136	0.001	-0.116	-0.026	-0.043	-0.042	-0.044	-0.132	-0.050	-0.072	-0.041	-0.046	-0.024	-0.039
	(0.168)	(0.107)	(0.178)	(0.106)	(0.105)	(0.104)	(0.144)	(0.132)	(0.127)	(0.105)	(0.106)	(0.104)	(0.117)	(0.143)	(0.128)
Time 3nd kid	-0.132	-0.009	-14.406	-0.327	-0.345	-0.344	-0.352	-0.150	-0.542	-0.376	-0.329	-0.399	-0.534	-0.569	-0.299
	(0.226)	(0.161)	(0.000)	(0.160)	(0.160)	(0.156)	(0.207)	(0.360)	(0.219)	(0.158)	(0.159)	(0.158)	(0.195)	(0.308)	(0.229)
Time 4th kid	-0.218	-0.521	0.038	-0.018	0.078	0.273	-5.067	-5.680	-12.044	0.147	0.170	0.002	1.106	-1.434	-2.344
	(0.433)	(0.276)	(0.000)	(0.256)	(0.253)	(0.246)	(0.916)	(0.908)	(0.000)	(0.244)	(0.249)	(0.246)	(0.382)	(0.648)	(0.638)

Table 13 (cont'd): Logit Coefficient of Best Response Probability for Married Female

(Standard Error in parenthesis; Choice 1 is the excluded class)

Spouse Choice	Choice 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	2.120	1.517	1.664	0.164	1.811	0.952	-0.013	1.611	0.628	0.369	1.539	1.392	0.039	2.087	0.775
	(0.264)	(0.223)	(0.939)	(0.323)	(0.345)	(0.299)	(0.366)	(0.523)	(0.623)	(0.312)	(0.434)	(0.342)	(0.366)	(0.615)	(0.786)
3	1.224	1.712	1.903	0.420	1.300	1.395	-0.003	1.110	1.061	0.551	1.432	1.395	0.366	1.119	1.370
	(0.198)	(0.147)	(0.743)	(0.195)	(0.249)	(0.188)	(0.214)	(0.421)	(0.397)	(0.196)	(0.310)	(0.232)	(0.223)	(0.529)	(0.530)
4	-9.888	2.112	-6.142	3.262	3.432	3.784	3.417	-11.592	3.577	3.378	3.663	3.949	4.718	3.710	4.771
	(1.078)	(1.134)	(1.315)	(1.124)	(1.154)	(1.111)	(1.108)	(1.135)	(1.247)	(1.120)	(1.205)	(1.137)	(1.084)	(1.543)	(1.314)
5	-10.517	1.996	5.948	2.406	4.893	3.885	3.018	5.511	3.911	3.366	5.164	4.051	3.788	3.833	4.984
	(0.916)	(1.027)	(1.323)	(1.019)	(0.985)	(0.986)	(0.991)	(1.044)	(1.152)	(0.971)	(1.006)	(1.018)	(0.941)	(1.462)	(1.195)
6	1.689	2.720	4.202	3.066	4.224	4.483	3.009	4.522	4.145	3.638	4.703	4.682	4.008	5.108	5.593
	(0.543)	(0.427)	(0.902)	(0.448)	(0.473)	(0.441)	(0.460)	(0.589)	(0.568)	(0.446)	(0.507)	(0.462)	(0.453)	(0.663)	(0.663)
7	0.081	0.407	3.095	-0.040	0.406	0.985	1.031	-10.111	1.782	0.922	0.754	1.251	1.076	2.274	2.736
	(0.826)	(0.517)	(0.990)	(0.601)	(0.660)	(0.537)	(0.625)	(0.600)	(0.685)	(0.585)	(0.738)	(0.593)	(0.669)	(0.906)	(0.852)
8	-10.054	1.744	3.784	1.065	2.941	2.167	-13.517	3.711	2.645	1.856	3.787	2.679	2.586	2.397	3.500
	(0.729)	(0.797)	(1.276)	(0.899)	(0.862)	(0.832)	(0.780)	(0.964)	(0.997)	(0.857)	(0.888)	(0.875)	(0.884)	(1.398)	(1.103)
9	0.712	1.190	3.321	1.495	2.686	3.203	1.889	3.331	3.061	2.691	3.732	4.033	3.093	4.343	5.177
	(0.489)	(0.338)	(0.857)	(0.368)	(0.394)	(0.350)	(0.391)	(0.526)	(0.502)	(0.358)	(0.431)	(0.375)	(0.364)	(0.609)	(0.599)
Constant	-1.843	-3.304	-18.480	-1.846	-4.269	-3.091	-6.185	-8.347	-10.911	-2.498	-3.544	-0.720	-6.742	-4.432	-2.633
	(0.928)	(0.680)	(3.105)	(0.956)	(1.060)	(0.833)	(1.440)	(1.940)	(1.944)	(0.946)	(1.195)	(0.926)	(1.333)	(1.998)	(1.606)
N	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834	26,834

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997.

Males' Ex-ante Best Response Probabilities

Table 14 presents the logit coefficient estimates of the ex-ante best response probabilities of a married male. Most of the effects of male's own variables on these probabilities are similar to that of single males. Table 12, however, shows that a male with a spouse who is college educated is less likely to choose not to work and provide high parental time investment. The same is true if his spouse is a high school graduate or attended some college. Apart from the effect of parental time investment in their 4th child, which reduces the likelihood of a male choosing not to work and provide low parental time investment, none of the other spouse characteristics has any effect on his choices.

The final panel in table 14 represents the reaction function of the male's choice probabilities to his spouse's choices. It shows that if the spouse chooses to work part time and not provide parental time investment or give birth (i.e. female's choice 2) then the male is less likely to choose choice 4, 5, and 9; that is he is less likely to work part time and provide high or low child care and less likely not to work and provide high time investment in children, and is more likely to choose to work full time and do nothing else. If the spouse chooses to work full time and give birth while not provide parental time investment the husband is least likely to choose not to work and provide low parental time investment. However, he is more likely to choose 5 or 7, which involve providing low parental investment while working full time or not working while providing high parental time investment. This is a case where the female is the main bread winner and gives birth, and the husband responds by providing the parental investment.

If the female choose to work part time while not giving birth, but provides low parental time investment, then the husband is more likely to choose choices 6 through 9; the first (i.e. male's choice 6) involves working full time while providing low parental time investment while the last three involve high parental time investment. A similar pattern holds for choice 7 (i.e. female choosing full time work, no birth, and low parental investment) except that there is a higher likelihood of choosing choices 3 and 4. If the female chooses choices 8 (i.e. not working, birth, and low parental time investment) then the male is least likely to choose 7 (i.e. not working and high parental time investment) and most likely to choose 4 (i.e. not working and low parental time investment). This highlights the fact that if the female does not work then the male has a higher probability of working. If the female chooses to work part time, give birth, and provides low parental time investment, then the husband has a higher likelihood of working in all possible combinations of parental time investment. On the other hand if the female chooses to work full time, give birth, and provide low parental time investment (i.e. choice 10) then the husband is more likely to provide the parental time investment (i.e. choices 4 through 9). This type of substitution pattern is highlighted through the other male's reactions to the female choices. Overall the reaction functions of both males and females display a certain degree of cooperation in their behavior. However, in cases in which females either do not give birth or provide no parental time investment, both spouses seem to focus on the maximizing labor income and leisure.

Table 14: Logit Coefficient of Best Response for Married Male

(Standard Error in parenthesis; Choice 1 is the excluded class)

Individual Variables	Choice 2	3	4	5	6	7	8	9
Black	0.297	-0.350	0.277	0.230	-0.368	-0.248	0.473	-0.330
	(0.332)	(0.288)	(0.675)	(0.684)	(0.313)	(0.737)	(0.559)	(0.325)
High Sch.	0.352	0.696	0.020	1.137	0.885	-0.219	0.905	0.864
	(0.280)	(0.218)	(0.459)	(0.649)	(0.241)	(0.401)	(0.483)	(0.256)
Some Col.	0.356	0.841	0.518	1.149	1.083	-0.414	0.734	1.067
	(0.314)	(0.243)	(0.525)	(0.682)	(0.266)	(0.463)	(0.522)	(0.280)
College	0.786	1.212	0.816	1.768	1.700	0.222	0.966	1.581
	(0.349)	(0.277)	(0.633)	(0.789)	(0.297)	(0.592)	(0.556)	(0.311)
Age	-0.316	-0.240	0.358	-0.694	-0.262	-0.122	0.057	-0.214
	(0.157)	(0.119)	(0.245)	(0.248)	(0.127)	(0.275)	(0.232)	(0.133)
Age Sq	0.004	0.002	-0.005	0.008	0.003	0.002	-0.000	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
No.of kids	-0.801	-0.102	1.665	1.019	1.144	2.331	2.835	1.122
	(0.494)	(0.412)	(1.213)	(0.819)	(0.444)	(0.879)	(0.869)	(0.458)
No.of kids Sq	0.170	-0.080	-0.668	-0.306	-0.385	-0.825	-0.769	-0.444
	(0.142)	(0.119)	(0.420)	(0.269)	(0.132)	(0.272)	(0.260)	(0.137)
Part work (t-1)	2.288	0.697	0.005	1.438	1.237	-0.077	1.023	1.510
	(0.377)	(0.275)	(0.453)	(0.677)	(0.410)	(0.469)	(0.556)	(0.405)
Part work (t-2)	-0.169	-0.740	-0.305	-0.347	-0.132	-0.630	0.212	-0.508
	(0.475)	(0.396)	(0.578)	(0.793)	(0.459)	(0.577)	(0.662)	(0.458)
Part work (t-3)	0.599	0.236	0.951	0.526	0.159	0.271	1.727	0.543
	(0.531)	(0.461)	(0.682)	(0.845)	(0.496)	(0.769)	(0.677)	(0.505)
Part work (t-4)	0.181	-0.547	-0.041	1.116	-0.457	-0.109	-0.891	-0.367
	(0.474)	(0.390)	(0.638)	(0.752)	(0.423)	(0.698)	(0.546)	(0.432)
Full work (t-1)	3.107	3.833	0.179	1.522	4.158	-0.414	1.642	3.912
	(0.372)	(0.264)	(0.502)	(0.701)	(0.380)	(0.496)	(0.597)	(0.382)
Full work (t-2)	-0.276	0.366	-1.027	0.039	0.863	-0.620	-0.103	0.323
	(0.469)	(0.396)	(0.660)	(0.765)	(0.442)	(0.594)	(0.729)	(0.445)
Full work (t-3)	-0.077	-0.091	-0.316	-0.393	-0.343	-0.022	0.145	-0.127
	(0.494)	(0.432)	(0.772)	(0.843)	(0.457)	(0.748)	(0.687)	(0.472)
Full work (t-4)	0.624	0.372	0.568	0.928	0.445	0.660	-0.660	0.311
	(0.434)	(0.364)	(0.598)	(0.774)	(0.383)	(0.694)	(0.520)	(0.395)
Age of 1st kid	-0.006	-0.021	-0.057	-0.130	-0.053	-0.028	-0.224	-0.081
	(0.039)	(0.025)	(0.056)	(0.099)	(0.027)	(0.047)	(0.087)	(0.028)
Age of 2nd kid	0.101	0.101	-0.078	0.184	0.119	-0.079	0.291	0.131
	(0.072)	(0.052)	(0.089)	(0.116)	(0.053)	(0.103)	(0.108)	(0.054)
Age of 3rd kid	-0.241	-0.237	-0.208	-0.507	-0.313	-0.186	-0.593	-0.301
	(0.120)	(0.077)	(0.110)	(0.161)	(0.080)	(0.122)	(0.331)	(0.084)
Age of 4th kid	0.122	0.205	-2.829	-0.221	0.012	0.212	-0.187	0.043
	(0.168)	(0.130)	(2.666)	(0.358)	(0.164)	(0.187)	(0.463)	(0.179)
Time 1st kid	-0.078	-0.047	0.156	0.304	0.188	0.250	0.334	0.329
	(0.081)	(0.063)	(0.096)	(0.102)	(0.063)	(0.088)	(0.097)	(0.064)
Time 2nd kid	-0.271	-0.146	-0.180	0.025	0.006	-0.131	0.055	0.112
	(0.115)	(0.081)	(0.149)	(0.135)	(0.082)	(0.147)	(0.122)	(0.083)
Time 3nd kid	0.937	0.703	0.354	0.723	0.863	1.387	0.771	1.097
	(0.307)	(0.277)	(0.551)	(0.511)	(0.280)	(0.416)	(0.410)	(0.283)
Time 4th kid	-0.287	-0.672	-2.131	1.134	-0.435	-0.188	-0.092	-0.489
	(0.397)	(0.328)	(1.166)	(0.688)	(0.334)	(0.470)	(0.660)	(0.354)
Female kids	0.105	0.094	0.555	0.314	0.123	0.271	0.291	0.027
	(0.213)	(0.160)	(0.264)	(0.304)	(0.162)	(0.262)	(0.252)	(0.164)
Spouse Variables								
High Sch.	-0.088	0.027	-0.805	-0.728	0.061	-1.430	-0.458	0.044
	(0.396)	(0.336)	(0.540)	(0.583)	(0.356)	(0.472)	(0.530)	(0.369)
Some Col.	-0.091	-0.048	-0.737	-1.163	0.094	-1.061	-0.665	0.069
	(0.425)	(0.361)	(0.586)	(0.635)	(0.381)	(0.538)	(0.580)	(0.394)
College	0.258	0.460	-1.378	-1.052	0.537	-1.855	0.032	0.431
	(0.464)	(0.394)	(0.855)	(0.753)	(0.415)	(0.781)	(0.627)	(0.428)
Age	-0.054	-0.186	-0.339	0.131	-0.151	-0.299	-0.261	-0.141
	(0.170)	(0.131)	(0.268)	(0.280)	(0.137)	(0.297)	(0.253)	(0.144)

3.4.2 Preference Parameter Estimates

Table 15 presents the GMM estimates of the parameters characterizing the utility of functions along with the various discount factors of the model. There are two sets of estimates; the first set consists of estimates of a baseline model where the parameters do not vary by demographic characteristics, and the second set consists of estimates of an extended model where the parameters vary by demographic characteristics and the education of the individuals in the households.

First, the top left hand panel of Table 15 shows that there are per-period utility costs of giving birth for females. This is demonstrated by the universal significant and negative coefficients associated with all choices in the per-period utility function that involve giving birth in the current period. This finding rationalizes the low frequency of these choices in the data and conforms to the finding of previous literature on fertility behavior (see Wolpin (1984)(84) and Hotz and Miller (1988)(42) for example).

While the utility for female is monotonically declining in the level of labor market work for no birth and low level of parental time (i.e. choices 5 through 7), this is not always the case for other choice permutations. This seems to be caused by the interaction of labor market choice with parental time investment; some levels of parental time investment seem to be preferred to no parental time if these choices do not involve low levels of leisure. This implies that there may be some level of consumption value to maternal time investment. For example, conditional on working part time in the labor market and not giving birth in the current period, the utility of mothers are increasing in the level of parental time investment. This monotonic relationship is not present conditional on working full time in the labor market and not giving birth in the current period. This may be due to the nonlinear nature of time requirements of jobs or occupations chosen by females. That is, the full time and part time classification does not fully capture the degree of effort or flexibility of hours associated with female job choices.

The top right hand panel of Table 15 presents the estimates for males. It shows that the disutility from working in nonlinear in the level of labor market work activities. Conditional on providing zero paternal time investment, males prefer working part time to either not

Table 14 (cont'd): Logit Coefficient of Best Response for Married Male

(Standard Error in parenthesis; Choice 1 is the excluded class)

Spouse Variables	Choice 2	3	4	5	6	7	8	9
Age Sq	0.000	0.002	0.004	-0.001	0.002	0.003	0.002	0.001
	(0.003)	(0.002)	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)	(0.002)
Part work (t-1)	0.232	-0.067	-0.581	0.971	0.102	0.510	0.356	0.306
	(0.377)	(0.328)	(0.705)	(0.576)	(0.337)	(0.634)	(0.565)	(0.342)
Part work (t-2)	-0.191	-0.503	-0.365	-0.123	-0.454	-0.427	-0.094	-0.407
	(0.387)	(0.329)	(0.669)	(0.564)	(0.337)	(0.736)	(0.550)	(0.343)
Part work (t-3)	0.003	0.259	0.947	0.233	0.256	-0.182	0.324	0.235
	(0.432)	(0.366)	(0.655)	(0.539)	(0.372)	(0.695)	(0.526)	(0.376)
Part work (t-4)	-0.346	-0.328	-0.949	-0.461	-0.393	-1.193	-0.518	-0.479
	(0.373)	(0.312)	(0.678)	(0.551)	(0.319)	(0.599)	(0.477)	(0.324)
Full work (t-1)	-0.312	-0.519	-1.116	0.512	-0.186	-0.394	0.844	-0.008
	(0.325)	(0.279)	(0.662)	(0.535)	(0.292)	(0.542)	(0.488)	(0.301)
Full work (t-2)	-0.030	-0.235	0.282	-0.157	-0.224	0.702	-0.204	-0.289
	(0.377)	(0.329)	(0.781)	(0.533)	(0.338)	(0.648)	(0.533)	(0.345)
Full work (t-3)	0.134	0.375	1.251	0.120	0.365	0.856	0.499	0.392
	(0.395)	(0.333)	(0.760)	(0.522)	(0.342)	(0.616)	(0.489)	(0.348)
Full work (t-4)	-0.192	-0.001	-0.536	-0.064	-0.065	-1.011	-0.926	-0.210
	(0.338)	(0.286)	(0.499)	(0.502)	(0.293)	(0.519)	(0.442)	(0.298)
Time 1st kid	0.152	0.122	0.032	0.113	0.092	0.028	0.068	0.062
	(0.085)	(0.065)	(0.092)	(0.123)	(0.066)	(0.102)	(0.091)	(0.067)
Time 2nd kid	-0.035	0.060	0.175	-0.172	-0.035	0.211	-0.178	-0.030
	(0.110)	(0.081)	(0.151)	(0.155)	(0.082)	(0.150)	(0.129)	(0.083)
Time 3nd kid	-0.027	0.148	0.596	0.390	0.287	0.142	0.355	0.205
	(0.182)	(0.124)	(0.269)	(0.339)	(0.131)	(0.310)	(0.416)	(0.136)
Time 4th kid	-0.113	0.004	-4.764	-0.757	0.191	0.137	0.686	0.277
	(0.318)	(0.238)	(1.313)	(0.545)	(0.262)	(0.417)	(0.401)	(0.290)
Spouse Choice								
2	0.798	0.708	-7.977	-7.524	-0.499	1.002	-7.642	-0.829
	(0.415)	(0.358)	(0.835)	(1.096)	(0.575)	(1.210)	(1.226)	(0.679)
3	0.354	0.802	-0.161	0.824	0.063	1.404	1.466	-0.461
	(0.304)	(0.261)	(0.984)	(1.178)	(0.339)	(1.055)	(1.229)	(0.364)
4	-0.386	0.473	-6.164	4.440	1.831	4.610	3.250	1.328
	(1.474)	(1.092)	(1.250)	(1.640)	(1.141)	(1.667)	(1.815)	(1.167)
5	-0.122	0.069	1.430	1.383	1.140	1.168	0.786	0.454
	(0.547)	(0.446)	(1.061)	(1.427)	(0.504)	(1.348)	(1.301)	(0.524)
6	1.504	1.492	2.418	2.769	2.865	3.323	3.109	2.186
	(0.854)	(0.781)	(1.679)	(1.567)	(0.815)	(1.456)	(1.470)	(0.825)
7	0.604	0.830	2.502	2.396	2.322	2.391	1.647	1.974
	(0.466)	(0.392)	(1.097)	(1.331)	(0.456)	(1.228)	(1.231)	(0.466)
8	-0.718	-0.460	2.423	2.116	1.060	-7.198	1.789	0.910
	(0.773)	(0.633)	(1.249)	(1.603)	(0.687)	(1.338)	(1.717)	(0.718)
9	6.506	6.774	-1.160	10.182	9.091	-1.574	9.375	8.853
	(0.673)	(0.401)	(0.951)	(1.423)	(0.448)	(1.432)	(1.656)	(0.462)
10	0.937	1.629	3.393	4.239	3.830	4.628	3.865	3.700
	(1.068)	(0.961)	(1.790)	(1.559)	(0.982)	(1.561)	(1.607)	(0.987)
11	-0.361	-0.081	0.906	2.506	1.384	2.199	2.126	1.345
	(0.535)	(0.432)	(1.078)	(1.429)	(0.490)	(1.213)	(1.290)	(0.500)
12	-0.075	0.547	-7.276	-7.612	2.158	1.873	2.971	2.224
	(0.847)	(0.649)	(1.177)	(1.456)	(0.693)	(1.706)	(1.452)	(0.702)
13	0.299	0.356	1.695	1.974	1.973	2.130	2.262	2.351
	(0.550)	(0.452)	(1.180)	(1.421)	(0.510)	(1.268)	(1.291)	(0.517)
14	0.395	0.478	3.171	3.066	2.587	4.139	3.675	2.531
	(0.924)	(0.875)	(1.168)	(1.632)	(0.901)	(1.387)	(1.539)	(0.905)
15	-0.447	-0.300	-6.222	-6.783	2.603	3.433	-6.367	2.656
	(1.172)	(0.967)	(1.136)	(1.438)	(0.994)	(1.630)	(1.462)	(1.001)
16	5.704	6.666	9.575	-1.741	9.070	9.779	10.682	9.833
	(1.060)	(0.391)	(1.431)	(1.197)	(0.453)	(1.669)	(1.376)	(0.444)
Constant	5.532	8.194	-3.548	4.329	3.242	4.242	-2.155	2.666
	(1.966)	(1.522)	(3.343)	(3.751)	(1.671)	(2.969)	(3.560)	(1.769)
N	16,548	16,548	16,548	16,548	16,548	16,548	16,548	16,548

Source: Data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID) between 1968 and 1997.

working or working full time. Males, however, prefer not working in the labor market to working full time in the labor market. A similar pattern holds conditional on providing low paternal time investment. This pattern, however, is reversed conditional on providing high paternal time investment. This seemingly counter intuitive finding, that males prefer some work to not working, is the way the model rationalizes the low proportion of males that not work in our data. Similar to females, there seems to be some level of consumption associated with paternal time investment in children.

The second panel of Table 15 presents the discount factors. It shows that the intergenerational discount factor (i.e. 0.20) is smaller than the intertemporal discount factor (i.e. 0.67). This implies that in the second to last period of their life, a parent value their child 20% of their own utility next period. The discount factor on the number child shows that the marginal increase in the value of the second child is 0.63 and of the third child is 0.48. Although the estimated discount factor of children is close to the literature, it cannot be compared directly to these estimates because other models do not include the life cycle dimension. For example, in our model, a parent with horizon of 10 years, discounts the consumption of an only child, for example, by an additional time discount β^{10} which is less that 0.2. Thus, without taking into account the time dimension involved in trade-offs parents make when they are young, these investments may seem to be consistent with a much lower discount factor on the children's utility.

The bottom panel of Table 15 presents the estimates of the utility from earnings and the per-period net cost of existing children. It shows that, as expected, utility is increasing with own earnings for both genders, irrespective of marital status. The coefficient on spouse earning for male is, however, negative and large in magnitude; this means that males utility declines in the earnings of their spouse. Since our model specification implies transferable utility between spouses in the game, these estimates imply that there is a transfer of utility to the spouse the higher the earnings of the spouse. This may also implies higher outside option for higher earning spouses. There is a similar effect for female however of a much lower magnitude. Finally, the bottom panel of Table 15 shows that for both married male and female there is a per-period net cost of existing children. However, there is a per-period net benefit from a single father; this may be because the fact that most children stay with their mother hence the fathers utility is higher when they are not living in the same household.

Table 15: GMM Estimates of Utility of Liesure and Discount Factors

(Standard Errors in Parenthesis; Choice 1 is the Excluded Class)

Utility of Leisure								
Female					Male			
Choice	Labor Market Work	Child Birth	Parental Time	(1)	Choice	Labor Market Work	Parental Time	(2)
2	Part time	None	None	-5.16	2	Part Time	None	0.24
				(4.0e-3)				(0.02)
3	Full Time	None	None	0.465	3	Full Time	None	-0.25
				(3.0e-3)				(0.02)
4	Full Time	Yes	None	-11.65	4	None	Low	-4.84
				(6.0e-3)				(0.02)
5	None	None	Low	-0.23	5	Part Time	Low	-4.06
				(0.02)				(0.01)
6	Part Time	None	Low	-0.64	6	Full Time	Low	0.61
				(0.01)				(0.01)
7	Full Time	None	Low	-1.12	7	None	High	-0.95
				(0.04)				(0.03)
8	None	Yes	Low	-1.85	8	Part Time	High	-2.12
				(0.01)				(0.01)
9	Part Time	Yes	Low	-4.87	9	Full Time	High	-0.28
				(0.01)				(0.01)
10	Full Time	Yes	Low	-2.73				
				(0.01)				
11	None	None	High	0.19				
				(0.02)				
12	Part Time	None	High	0.2				
				(0.02)				
13	Full Time	None	High	-0.28				
				(0.02)				
14	None	Yes	High	-0.12				
				(0.01)				
15	Part Time	Yes	High	-2.78				
				(0.01)				
16	Full Time	Yes	High	-3.19				
				(0.01)				
Discount Factors								
Intertemporal			β	0.67				
				(0.02)				
Intergenerational			λ	0.20				
				(0.02)				
Number Children			ν	0.65				
				(0.01)				
Utility of Earnings and Net Cost of Children								
Married own earnings				0.65	Married own earnings			0.50
				(0.001)				(0.005)
Married Spouse earnings				0.05	Married Spouse earnings			-0.38
				(0.001)				(0.002)
Married number of children				-0.12	Married number of children			-0.69
				(0.02)				(0.002)
Single earnings				0.64	Single earnings			-0.19
				(0.004)				(0.004)
Single number of children				-0.04	Single number of children			0.39
				(0.006)				(0.004)
N				50,514				

3.5 MEASURING THE QUALITY-QUANTITY TRADE-OFFS AND THE RETURN TO PARENTAL INVESTMENT

The dynastic model provide a nature measure of the quality-quantity trade-offs and the returns to parental investment. Lets consider an parent entering the final period of his/her life and lets further assume for convenient that he/she has completed fertility decisions. This assumption is without lost of generality because we assume that females can not child after the age of 45 in our implementation and so this in more relevant for male who are significantly older than their spouse. Taking the expectation over the choices of the last term in equation (3.7) we can write the expected value of children at age T as

$$\overline{V}_{N\sigma}(x_T) = \sum_i \left(p_{-\sigma iT}(k_{-\sigma iT}|x_T) \left[\sum_j p_{\sigma jT}(k_{\sigma jT}|k_{-\sigma iT}, x_t) \overline{V}_{N\sigma}(k_{jiT}; x_T) \right] \right) \tag{3.19}$$

The average quality of a child is given by $\frac{N_T^{1-v}\overline{V}_{N\sigma}(x_T)}{N_T}$, we can therefore measure the quality-quantity trade-off as

$$\begin{aligned} \Lambda_{N\sigma}(x_t) &\equiv \frac{\partial \log\left(\frac{N_T^{1-v}\overline{V}_{N\sigma}(x_T)}{N_T}\right)}{\partial N_T} && (3.20) \\ &= \left[1 - v + \frac{\partial\left(\frac{\overline{V}_{N\sigma}(x_T)}{N_T}\right)}{\partial N_T} \frac{N_T}{\left(\frac{\overline{V}_{N\sigma}(x_T)}{N_T}\right)} \right] \frac{1}{N_T} && (3.21) \end{aligned}$$

This measure of quantity-quality trade-off has two components: the first element in Equation 3.20 , $1 - v$, reflects the rate of increase in utility with an additional child, and the elasticity component reflect the rate of decline in the average quality per child. The model then exhibits a quality-quantity trade-off if the elasticity of the average quality of a child is larger (in magnitude) than the rate of the increase in parental utility. In general, this may not hold in equilibrium because, as noted in Hill and Stafford (1974), when parents make the time allocation decisions in children they take into account the differential effect of time on the different children which affect this trade-off. Next, we measure the return to parental time investment as

$$
\begin{aligned}
\Lambda_{D\sigma}(x_t) &\equiv \frac{\partial \log \left(\frac{N_T^{1-^{1-v}} \overline{V}_{N\sigma}(x_T)}{N_T} \right)}{\partial D_T} \\
&= \left[\frac{\partial \overline{V}_{N\sigma}(x_T)}{\partial D_T} \right] \frac{1}{\overline{V}_{N\sigma}(x_T)}.
\end{aligned}
\tag{3.22}
$$

This measures the aggregated return to parental time investment which measures the impact of parental time input on educational attainment of children, their skills and therefore life time earnings, as well as their marriage market outcomes and life time choices. If a parent provides an additional unit of time investment, each child in the household receives an equal share of the time. Thus, the above measure depends on the number of children in the household.

The valuation function of the next generation (from the entire stock of children) $\overline{V}_{N\sigma}(x_T)$, is calculated by using the estimated structural parameters to simulate the model for each individual in our data and calculate their terminal valuation as age 55. Table 16 presents the estimates of the quality-quantity trade-off and these aggregate return to parental time investment. The standard errors are model errors which account for the variation in the outcome of the model prediction as well as estimation errors.

3.5.1 Quality-Quantity Trade-offs

The coefficients on the number of children in Table 16 measure the quality-quantity trade-offs; the coefficients on the linear term show that there is a trade-off for both black and white individuals, that is $1 - v < -\left(\frac{\partial \left(\frac{\overline{V}_{N\sigma}(x_T)}{N_T} \right)}{\partial N_T} \frac{N_T}{\left(\frac{\overline{V}_{N\sigma}(x_T)}{N_T} \right)} \right)$. The coefficients on the quadratic term shows that effect is nonlinear in nature, which means that parents are not employing a nondiscriminatory time allocation strategy (see Hanuschek (1992)(38) for a similar finding). By comparing the estimates across race, we see that the quality-quantity trade-offs for black are similiar to that for whites and an increase in number of children implies a reduction in the average valuation function of each child. Note that we find that the fertility behavior of married couples does not vary significantly with race. Turning to the gender of the children;

we find that the quality-quantity trade-off is significantly less for female children. This also varies significantly by race, with whites having more concave quality-quantity trade-off for female children than blacks.

In summary, we find significant quality-quantity trade-off in our model. The quantity quality trade-off is smaller the larger the education of the fathers (with fathers education reducing the trade-off more than mothers education) implying a lower reduction in the average quality of a child as a result of increase in the number of children if the father is a white male. The effect of edcuation is insignificant for black fathers. Also the education of the mother does not matter in terms of changing the effect of quality quantity trade off for a given number of children. This is related to the fact that mother's time investment work throught the production function where college educated mother increases the likelihood of her children attend college.

This result suggests that the lower fertility of more educated, high income households is driven by the high cost of time of educated parents. We also find that female have higher valuation functions (i.e. female child expected lifetime utility is higher than a male child), this is despite the fact that there is a female "tax" in the labor market. However, despite lower earnings, females are more likely to obtain higher education given equal inputs and education is highly compensated in the labor market. However, given education level, the valuation function of females are higher because they receive high utility from their husband's income because there are endowed with the ability to bear children and males place significant value on children.

3.5.2 The Return to Parental Time Investment

The coefficients on the parental time investment in Table 16 summarize our estimates of the return to parental time investment. They show that maternal parental time investment has a significantly higher return than parental time investment for whites; the estimated elasticity of father's time investment is about 40% of that of mother's time investment. For blacks, the maternal time investment is insignificantly estimated implying a no direct effect of mother's time, whereas black father's time investment matters and the effect depends on the children composition in terms of gender.

Despite the fact that we found no clear patterns suggesting that mothers' time is more valuable than father's time in terms of the education production function, white mothers have a higher return. One possibility is that the interaction of spouses within households is the cause. For example, if there are increasing return to scale to a parent's investment, and if because of the gender "tax" in the labor market, mother provides more parental time, this could explain the higher return to maternal time investment. However, there is nothing in the specification of our model that allow for increasing return to scale in the education production or skill function. Therefore this result is driven by the differential impact of maternal and parental time on the education outcomes of children. The estimates of the education production function Table 6 show that paternal time increases the probability of graduating from high school and getting some college while mother‘s time increases the probability of having a college degree. Thus paternal time truncates bad outcomes (i.e., not graduating from high school) while maternal time investment increases the probability of being a high achiever. Our estimates reveal that maternal time has a higher impact overall than paternal time because of the higher return of graduating college in both the labor and marriage markets. This result illustrates the advantage of aggregating the different outcomes of children when measuring the returns to parental time investment.

Turning to race, we find that the return to maternal time investment is insignificantly estimated (if not negative, when we look at the sign of the coefficient estimate) for blacks. However the effect of maternal time investment of a white mother is significant and positive. Moreover this effect is independent of the number of children and the gender of the children. Basically a white female who increases her time with children increases her returns by doing so, however given the quality-quantity trade-off she faces, the aggregate return to parental time investment is decreasing.

For the paternal time investment, the results are more complicated and depend on both the number and gender of the children. For black fathers, spending time increases the returns for the first child, and the increase is higher if the child is a girl. However for the additional male children, the total effect of time investment is negative, meaning decreasing the aggregate return. For additional female children, the reverse is true; the returns are

Table 16: OLS Estimates of Aggregated Return to Parental Time Investment

Dependent Variable: $\text{Log}(\frac{(N_{\sigma T})^{1-v}}{N_{\sigma T}}\overline{V}_{N\sigma}(x_T))$

(Standard Errors in Parenthesis)

	Baseline Model	
Variables	Black	White
Number Children	-0.663	-0.695
	(0.003)	(0.005)
Number Children Squared	0.067	0.070
	(0.0005)	(0.001)
Number of Female Children	0.112	0.187
	(0.001)	(0.002)
Number of Female Children Squared	-0.014	-0.023
	(0.0002)	(0.0004)
Mother: High School	-0.0004	-0.001
	(0.001)	(0.002)
Mother: Some College	0.002	-0.001
	(0.001)	(0.002)
Mother: College	-0.0001	-0.0002
	(0.001)	(0.002)
Father: High School	-0.001	0.003
	(0.001)	(0.002)
Father : Some College	-0.0007	0.004
	(0.001)	(0.002)
Father : College	-0.0003	0.006
	(0.001)	(0.002)
Mother's Time Investment	-0.0005	0.004
	(0.0005)	(0.001)
x Number of Children	-0.00008	0.0001
	(0.0002)	(0.003)
x Number Female Children	0.0002	-0.0001
	(0.0001)	(0.0001)
Father's Time Investment	0.0005	0.0015
	(0.0005)	(0.001)
x Number of Children	-0.0004	-0.001
	(0.0001)	(0.0003)
x Number Female Children	0.0009	0.002
	(0.0001)	(0.0001)
Constant	1.401	1.505
	(0.005)	(0.08)
N	6,720	6,720
R-squared	0.989	0.978

increasing in number of female children. However this increase is much less than the effect of quality-quantity trade-off. Therefore the final effect of having more children and increasing the time investment will end up as decreasing the aggregate return measure given in Table 16.

For the white fathers, time investment increases the returns for the first two children for both male and female children, although having a girl increases it more. After the second child, the effect turns to negative for additional male children, whereas increases with the additional female children. Similar to the black fathers, the final effect of having more children and increasing the time investment will decrease the aggregate return measure.

The return to time investment of black mother is estimated as insignificant. One explanation for this result could be the family structure differences between blacks and whites. Black provides lower maternal time investment. There is a significantly higher number of black single mothers than white single mothers, and single mothers invest less in their children because it is more costly for them to specialize in parental investment.

Table 16 also shows that maternal time investment is important for white females while not for black females. There are no differences in the return to maternal time investment between boys and girls for white mothers. However for fathers (both black and white), the returns to paternal time investment are significantly higher for girls. This may suggest that fathers act in a achievement maximizing manner, favoring high ability children in the family. Since girls have a higher likelihood of high education outcome than boys. This findings conflict with the finding in Hanuschek (1992)(38) that parents seem to act in compensatory or neutral manner. Our results hold for both blacks and whites while the results in Hanuscek (1992)(38) were restricted to blacks.

3.6 CONCLUSION

In this paper we developed and estimated a model of dynastic households in which altruistic individuals choose fertility, labor supply, and time investment in children sequentially, using data on two generations from the PSID. We then use the estimates to quantify the quality-quantity trade-offs and the return to parental time investment in children. Our preliminary analysis shows that parental investment in children varies significantly across gender, race, education levels, and the household composition. It also shows that after controlling for gender, education levels, and household composition, the differences across race are significantly reduced.

The structural estimates show that there are significant transfers between spouses within households and that females with higher earnings potentials receive larger transfers. The production function estimates show that both maternal and paternal time investment increase the likelihood of higher educational outcome of their children. However, the impact is complementary; fathers' time investment increases the probability of graduating from high school and getting some college education while mothers' time increases the probability of achieving a college degree. The estimates of the education production-function show that girls have a higher likelihood than boys of achieving higher education levels, and that blacks have higher variance than white in their educational outcomes, after controlling for parental inputs. Specifically, blacks have higher a higher probability of not completing high school as well graduating from college than whites.

We find that the intergenerational discount factor (i.e. 0.20)and the intertemporal discount factor is 0.67. This implies that in the second to last period of their life, a parent value their child 20% of their own utility next period. The discount factor on the number child shows that the marginal increase in the value from the second child is 0.63 and from the third child is 0.49. Although the estimated discount factor of children is significantly larger than previous estimates in the literature, it cannot be compared directly to these estimates because other models do not include life cycle. For example, in our model, a parent with horizon of 10 years, discounts the consumption of an only child by an additional time discount β^{10} which is less that 0.2. Thus, without taking into account the time dimension involved in trade-offs parents make when they are young, these investments may seem to be consistent with a much lower discount factor on the children's utility.

We find significant quality-quantity trade-off in our model. This trade-off is measured in terms of the rate of increase in utility of parents and the rate of the decline in the average life time utility per child resulting from having an additional child. The level of investment per child is smaller the larger the number of children, thus, this decline in the per child investment is driven by the time constraint and the opportunity costs of time and not by the properties of the production function technology of children. The negative relationship between income (education) and fertility is therefore explained by the higher opportunity cost of time of educated parents in terms of forgone earnings. We find similar quality-quantity trade-off for blacks and whites. However the reasons for the same level of quality-quantity trade-off could be completely different. For instance the returns to time investment for black females are estimated as insignificant. Therefore the effect of time investment does not seem to have a direct link for black mothers, but interacting with the other model features.

Interestingly, we find that females have higher valuation functions (i.e. female child value is higher than that of a male child). Despite the fact that females earn less than men with the same productive characteristics, females are more likely to obtain a higher education level than males, given equal amount of parental inputs and education is highly compensated in the labor market. However, even given education level the valuation function of females are higher than males. Despite the fact that females earn less than men with the same productive characteristics, females are more likely to obtain a higher level of education than males, given equal amount of parental inputs and education is highly compensated in the labor market. However, even given the same levels of education the valuation function of females are higher than males because they receive significant transfers from their husband's income. These findings can be rationalized by the fact than females are endowed with birth decisions and males value children, but cannot make decisions to have them.

We find that the overall returns to fathers' time investment is only 40% that of mothers' time investment for white individuals. Maternal time investment increases the probability of a child graduating from college, and a college degree increases the returns in both the labor and the marriage markets. There are race differences in the returns to paternal time investment, and this interacts with the composition of the children in terms of gender in the household. Maternal time investment of black mothers are estimated as insignificant This may be related to different family structures between black and whites and the way this affects the channel how the time investment works.

Finally, the returns to maternal time investment is independent of the gender of the child, however the father's time favors girls more. This implies that fathers act in a achievement maximizing manner, favoring high ability children in the family. Since girls already have a higher likelihood of achieving a high level of education than boys, father seems to invest more time in girls than in boys as the number of children increases.

4.0 BIBLIOGRAPHY

[1] Aguirregabiria, Victor and Pedro Mira, “Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models,” 2002, *Econometrica*, Vol. 70, pp. 1519-1543.

[2] Albanesi, Stefania and Claudia Olivetti, “Maternal Health and the Baby Boom,” 2010, NBER Working Paper 16146.

[3] Altug, Sumru and Robert. A. Miller, “Household choices in equilibrium,” 1990, *Econometrica*, Vol. 58, pp. 543-570.

[4] Altug, Sumru and Robert. A. Miller, “The Effect of Work Experience on Female Wages and Labour Supply,” 1998, *Review of Economic Studies* Vol. 65, pp. 45–85.

[5] Alvarez, Fernando, "Social Mobility, "The Barro–Becker Children Meet the Laitner–Loury Dynasties," 1999, *Review of Economic Dynamics*, Vol. 2(1), pp. 65-103.

[6] Aughinbaugh, Alison, "Reapplication and Extension: Intergenerational Mobility in the United States," 2000, *Labor Economics*, Vol. 7, pp. 785-796.

[7] Gary S. Becker, " Theory of the Allocation of Time," 1965, *The Economic Journal*, Vol. 75, No. 299, pp. 493-517

[8] Barro, Robert J.and Gary S. Becker, "A Reformulation of the Economic Theory of Fertility," 1988, *Quarterly Journal of Economics* Vol.103(1), pp. 1-25.

[9] Barro, Robert J. and Gary S. Becker, "Fertility Choice in a Model of Economic Growth," 1989, *Econometrica*, Vol. 57(2), pp. 481-501

[10] Barro, Robert J.and Gary S. Becker, "A Reformulation of the Economic Theory of Fertility," 1994, *American Economic Review* Vol.84, pp. 396-418.

[11] Becker, Gary S. and Michael Grossman and Kevin M. Murphy, "An Empirical Analysis of Cigarette Addiction," 1989, *Econometrica*, Vol. 57(2), pp. 481-501

[12] Becker, Gary S. and Nigel Tomes, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", 1979, *The Journal of Political Economy,* Vol. 87, pp.1153-1189.

[13] Becker, Gary S. and Nigel Tomes, "Human Capital and the Rise and Fall of Families" *Journal of Labor Economics,* 1986 vol. Vol. 4(3), pp. 1-39.

[14] Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy,* 1967 vol. Vol. 75, pp. 352-365.

[15] Berhman, J., A. Foster, M. Rosenzweig and P. Vashishtha, " Women's Schooling, Home Teaching and Economic Growth", 1999, *The Journal of Political Economy,* Vol. 107(4), pp 682-714.

[16] Black, Sandra E. and Paul J. Devereux "Recent Developments in Intergenerational Mobility," 2011 *Handbook of Labor Economics,* Vol. 4b.

[17] Blau, Francine D. and Lawrence M. Kahn, "Gender Differences in Pay," 2000, *Journal of Economic Perspectives,* Vol.14, pp. 75-99.

[18] Chamberlain, Gary, "Asymptotic Efficiency in Semiparametric Models with Censoring",1986, *Journal of Econometrics*, Vol. 32(2), 189–218.

[19] Chiswick, Barry R., "Differences in Education and Earnings Across Racial and Ethnic Groups: Tastes, Discrimination, and Investments in Child Quality," 1988, *Quarterly Journal of Economics,* Volume103(3), pp. 571-97.

[20] Zhiqi Chen, Frances Woolley, "A Cournot–Nash Model of Family Decision Making," 2001, *The Economic Journal,* Vol. 111, Issue 474, pages 722–748

[21] Pierre-André Chiappori, "Rational Household Labor Supply," 1988, *Econometrica,* Vol. 56, No. 1, pp. 63-90

[22] Pierre-André Chiappori and Donni Olivier, "Non-Unitary Models of Household Behavior: A Survey of the Literature," 2009, IZA Discussion Paper no. 4603

[23] Flavio Chnha and James Heckman, "The Technology of Skill Formation," 2007, NBER Working paper # 12840

[24] Datcher-Loury, Linda, "Effects of mother's home time on children's schooling" 1988, *The Review of Economics and Statistics*, Vol. 70(3), pp. 367-373.

[25] Del Boca, Daniela and Christopher J. Flinn, "Rationalizing Child-Support Decisions," 1995, *The American Economic Review,* Vol. 85, No. 5, pp. 1241-1262

[26] Del Boca, Daniela and Christopher J. Flinn, " Endogenous Household Interaction," 2010, Working Paper, New York University.

[27] Del Boca, Daniela, Christopher J. Flinn and Matthew Wiswall, "Household Choices and Child Development," 2010, Working Paper, New York University.

[28] Doepke, Matthias, "Child mortality and fertility decline: Does the Barro-Becker model fit the facts?", 2004, *Journal of Population Economics*, Vol. 18, pp. 337-366.

[29] Doepke, Matthias and Michele Tertilt, "Women's Liberation: What's in it for Men?", 2009, *Quarterly Journal of Economics*, Vol 124 (4): pp. 1541-1591.

[30] Echevarria, Cristina and Antonio Merlo, "Gender Differences in Education in a Dynamic Household Bargaining Model", 1999, *International Economics Review*, Vol. 40, pp. 265-286.

[31] Eckstein, Zvi and Kenneth I. Wolpin, "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models", 1989, *Journal of Human Resources*, Vol. 24, pp. 562-598.

[32] Gayle, George-Levi and Christelle Viauroux, "Root-N Consistent Semi-parametric Estimators of a Dynamic Panel Sample Selection Model", 2007, *Journal of Econometrics*, Vol. 141, pp. 179-212.

[33] Gayle, George-Levi and Limor Golan, “Estimating a Dynamic Adverse Selection Model: Labor Force Experience and the Changing Gender Earnings Gap”, Forthcoming, *Review of Economic Studies*.

[34] Gayle, George-Levi and Robert A Miller, "Life-Cycle Fertility and Human Capital Accumulation", 2006, Tepper School of Business, Carnegie Mellon University.

[35] Gayle, George-Levi, Limor Golan and Mehmet Ali Soytas, “Estimation of Intergenerational Lifecycle Discrete Choice Models”, 2010, Tepper School of Business, Carnegie Mellon Unviersity.

[36] Goldin, Claudia, "Understanding the Gender Gap: An Economic History of American Women," 1990, Oxford University Press.

[37] Guryan, Jonathan, Erik Hurst and Melissa Kearney,"Parental Education and Parental Time With Children," 2008, J*ournal of Economic Perspectives*,Vol. 22(3), pp. 23-46.

[38] Hanuscek, Eric a. : " The Trade-off between Child Quantity and Quality", 1992, *The Journal of Political Economy*, Vol. 100 (1), pp. 84-117.

[39] Haveman, Robert and Barbara Wolf, "The Determinants of Children’s Attainments: A Review of Methods and Findings," 1995, *Journal of Economic Literature*, Vol. 33, pp. 1829-1878.

[40] Hill, C. Russell and Frank P. Stafford, "Parental Care of Children: Time Diary Estimates of Quantity, Predictability, and Variety," 1980, *The Journal of Human Resources*, Vol. 15(2), pp. 219-239.

[41] Hotz, V. Joseph and F.E. Kydland and G.L. Sedlacek, "Intertemporal substitution and labor supply", 1988, Econometrica, Vol. 56, pp. 335-360.

[42] Hotz, V. Joseph and Robert A. Miller, "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply", 1988, *Econometrica*, Vol. 56, pp. 91-118.

[43] Hotz, V. Joseph and Robert A. Miller, "Conditional choice probabilities and the estimation of dynamic models", 1993, *The Review of Economic Studies*, Vol. 60, pp. 497-530.

[44] Hotz, V. Joseph and Robert A. Miller and S. Sanders and J. Smith, "A simulation estimator for dynamic models of discrete choice", 1994, The Review of Economic Studies, Vol. 61, pp. 265-289.

[45] Houtenville, Andrew J. and Karen Smith Conway "Parental Effort, School Resources, and Student Achievement," *The Journal of Human Resources,* 2008 Vol. 43(2), pp 437-453.

[46] Jones, Larry E. and Alice Schoonbroodt and Michele Tertilt, "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?" 2008, Working Paper.

[47] Juster, F. Thomas and Frank P. Stafford, "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement" 1991, *Journal of Economic Literature*, Vol. 29(2), pp. 471-522.

[48] Kang, Ari C., "Parental Choices and the Labor Market Outcome of Children", 2009, Tepper School of Business, Carnegie Mellon Univerisity.

[49] Kooreman, Peter and Arie Kapteyn, "A Disaggregated Analysis of the Allocation of Time within the Household," 1987, *The Journal of Political Economy* Vol. 95(2), pp. 223-249.

[50] Laitner, John, "Random earnings differences, lifetime liquidity constraints, and altruistic intergenerational transfers," 1992, *Journal of Economic Theory*, Vol. 58(2), pp. 135-170.

[51] Leibowitz, Arleeen "Home Investments in Children," 1974, *The Journal of Political Economy*, Vol. 82(2), pp. 111-131.

[52] Leibowitz, Arleeen "Parental inputs and children's achievement," 1977, *The Journal of Human Resources*, Vol. 12(2), pp. 242-251.

[53] Loury, Glenn C., "Intergenerational Transfers and the Distribution of Earnings," 1981, *Econometrica*, Vol. 49(4), pp. 843-867.

[54] Lundberg, Shelly and Robert A. Pollak., "Separate Spheres Bargaining and the Marriage Market," 1993, *Journal of Political Econom*y, Vol. 101, pp. 988-1010.

[55] MaCurdy, Thomas. E., "An Empirical Model of Labor Supply in a Life-Cycle Setting," 1981, *Journal of Political Economy*, Vol. 89, pp. 1059-1085.

[56] Manser, Marilyn and Murray Brown, "Marriage and Household Decision-Making: A Bargaining Analysis," 1980, *International Economic Review,* Vol. 21, No. 1, pp. 31-44

[57] McElroy Marjorie B. and Mary Jean Horney, "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand," 1981, *International Economic Review*, Vol. 22, No. 2, pp. 333-49

[58] Mazzocco, Maurizio, "Household Intertemporal Behavior: A Collective Characterization and a Test of Commitment," 2006, UCLA Working Paper.

[59] Milgrom, P., and J. Roberts, "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities", 1990, *Econometrica*, Vol. 58(2), pp. 1255–1277.

[60] Milgrom, P., and C. Shannon, "Monotone Comparative Statics," 1994, *Econometrica*, Vol. 62, pp. 157–180.

[61] Miller, A. Robert, "Job Matching and Occupational Choice," 1984, *Journal of Political Economy*, Vol. 92, pp. 1086-1120.

[62] Miller, A. Robert, "Estimating Models of Dynamic Optimization with Microeconomic Data," 1997, *Handbook of Applied Econometrics: Microeconomics (Vol 2), ed. by H. Pesaran and P. Smidth. Oxford, U.K.: Blackwell*, pp. 246-299.

[63] Mulligan, Casey, "Parental Priorities and Economic Inequality", The University of Chicago Press, 1997.

[64] Murnane, Richard J., Rebecca A. Maynard, and James C. Ohls, "Home resources and children's achievement", 1981, T*he Review of Economics and Statistics*, Vol. 63(3), pp. 369-377.

[65] Neal, Derek, "Why Has Black–White Skill Convergence Stopped?," 2006, Handbook of the Economics of Education Volume 1, pp. 511-76.

[66] Newey, Whitney K., "The Asymptotic Variance of Semiparametric Estimators," 1994, *Econometrica*, Vol. 16, pp. 1–32.

[67] Nijman, T. and M. Verbeek, "Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function", 1992, *Journal of Applied Econometric*s, Vol. 7, pp. 243–257.

[68] Pakes, Ariel,"Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," 1996, *Econometrica*, Vol. 54, pp. 755-784.

[69] Pakes, Ariel, "Dynamic Structural Models, Problems and Prospects: Mixed Continuous Discrete Controls and Market Interactions," 1994, Advances in Econometrics, Proceedings of Sixth World Congress of Econometric Society, ed. by C. Sims Cambridge: Cambridge University Press, pp 171-259.

[70] Rios-Rull, Jose-Victor and Virginia Sanchez-Marcos, "College Attainment of Women," 2002, *Review of Economic Dynamics*, Vol. 5, pp. 965-998.

[71] Rosenzweig, Mark R.; Wolpin, Kenneth I., "Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement" *Journal if Human Resources* 1994, Vol. 29(2), pp 670-693.

[72] Rupert, Peter, Rogerson Richard and Wright, Randall, "Homework in labor economics:

Household production and intertemporal substitution", 2000, *Journal of Monetary Economics*, Volume 46(3), pp. 557-579.

[73] Rust, John "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," 1987, *Econometrica,* Vol.55, pp 999-1033.

[74] Rust, John, "Maximum Likelihood Estimation of Discrete Control Processes," 1988, *SIAM Journal of Control and Optimization,* Vol.26, pp 1006-1024.

[75] Rust, John, "Estimation of Dynamic Structural Models, Problems and Prospects: Discrete Decision Processes," 1994a, *Advances in Econometrics, Proceedings of the Sixth World Congress of the Econometric Society, ed. by C. Sims. Cambridge: Cambridge University Press,* pp 119-170.

[76] Rust, John, "Structural Estimation of Markov Decision Processes," 1994b, *Handbook of Econometrics, Volume 4, ed. by R. E. Engle and D. McFadden. Amsterdam: North-Holland, Vol.26*, pp 3082-3139.

[77] Rust, John, "Numerical Dynamical Programming in Economics," 1996, *Handbook of Computational Economics, ed. by H. Amman, D. Kendrick, and J. Rust. Amsterdam: Elsevier-North Holland, Chapter 14..*

[78] Smith, James P. and Michael Ward, "Women in the Labor Market and in the Family," 1989, *Journal of Economic Perspectives*, Volume 3, pp. 9-23.

[79] Solon, Gary and Mary Corcoran and Roger Gordon and Deborah Laren, "Sibling and Intergenerational Correlations in Welfare Program Participation," 1988, *Journal of Human Resources*, Volume 23, 388-396.

[80] Solon, Gary, "Intergenerational Income Mobility in the United States," 1992, *American Economic Review*, Volume 82, 393-408.

[81] Todd, E. Petra and Kenneth D. Wolpin, "On the Specification and Estimation of the

Production Function for Cognitive Achievement," 2003, *The Economic Journal,* Vol.113, pp 1–218.

[82] Topkis, D., *Supermodularity and Complementarity.* Princeton University Press, Princeton, NJ.

[83] Watanabe, Takahiro and Hideaki Yamashita"Existence of a Pure Strategy Equilibrium in Markov Games with Strategic Complementarities for Finite Actions and States", 2010.

[84] Wolpin, Kenneth I.,"An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," 1984, *Journal of Political Economy*, Vol. 92, pp. 852-874.

[85] Wolpin, Kenneth I.,"Public-Policy Uses of Discrete-Choice Dynamic Programming Models," 1996, *American Economic Review*, Vol. 86, pp. 427-432.

[86] Zabel, J.,"Estimating Fixed and Random Effects Model with Selectivity", 1992, *Economics Letters*, Vol. 40, pp. 269–272.

APPENDIX A

APPENDIX FOR CHAPTER 2

In this appendix we will show the necessary definitions that will lead us to the representation in equation (2.8).

The period T conditional valuation function is given as follows in equation (2.6):

$$v(k_{jT};x_T) = u(k_{jT},x_T) + \lambda \frac{(N_T+b_T)^{1-v}}{(N_T+b_T)} \overline{V}_N(k_{jT};x_T)$$

and since the state vector x_T includes $D_1,..,D_T$, $H(x_0'|x_T,k_{jT})$ is obtained as the expression $(N_T + b_T)^{1-v} \frac{\overline{V}_N(k_{jT};x_T)}{(N_T+b_T)}$ at the corresponding state x_T. Therefore:

$$v(k_{jT};x_T) = u(k_{jT},x_T) + \lambda \sum_{x_0} V(x_0')H(x_0'|x_T,k_{jT})$$

The period $T-1$ conditional valuation function is given as follows using equation (2.4)

$$v(k_{jT-1};x_{T-1}) = u(k_{jT-1},x_{T-1}) + \beta \sum_{x_T} V(x_T)F(x_T|x_{T-1},k_{jT-1}).$$

replacing for $V(x_T)$:

$$\begin{aligned} v(k_{jT-1};x_{T-1}) &= u(k_{jT-1},x_{T-1}) \\ &+\beta \sum_{x_T} \left(\sum_{k_T} [u(k_T,x_T) + E_\varepsilon(\varepsilon_T|k_T) + \lambda \sum_{x_0} V(x_0')H(x_0'|x_T,k_T)]p(k_T|x_T) \right) \\ &\times F(x_T|x_{T-1},k_{jT-1}) \end{aligned}$$

which is equal to:

$$
\begin{aligned}
\upsilon(k_{jT-1};x_{T-1}) &= u(k_{jT-1},x_{T-1}) \\
&+\beta\sum_{x_T}\left(\sum_{k_T}[u(k_T,x_T)+E_\varepsilon(\varepsilon_T|k_T)]p(k_T|x_T)\right)F(x_T|x_{T-1},k_{jT-1}) \\
&+\beta\lambda\sum_{x_T}\left(\sum_{k_T}\left[\sum_{x_0}V(x_0')H(x_0'|x_T,k_T)\right]p(k_T|x_T)\right)F(x_T|x_{T-1},k_{jT-1})
\end{aligned} \tag{A.1}
$$

In order to get the representation in (2.8), we need to define the $s-t$ transitions, assuming the state vector has a finite domain: $x_0 \in X = \{x_0^1, ...x_0^L\}$ and $x_t \in X_t = \{x_t^1, ...x_t^{L_t}\}$ for $t \in \{1, 2, ...T\}$.

Definition 2. *Let $\bar{F}^{k_t}_{t+1|t}$ be the $L_t \times L_{t+1}$ matrix of conditional transition matrix for the state from period t to period $t+1$ given the choice k_t in period t. Let $F^{k_t}_{t+1|t}(x^m_{t+1}|x^l_t)$ denote the mth column and lth row of the matrix $\bar{F}^{k_t}_{t+1|t}$ and the $L_{t+1}\times 1$ vector $\bar{F}_{k_t}(x_{t+1}|x_t)$ is defined as: $\bar{F}_{k_t}(x_{t+1}|x_t) = (F_{k_t}(x^1_{t+1}|x_t), ..., F_{k_t}(x^{L_{t+1}}_{t+1}|x_t))'$. Define $L_t \times L_{t+1}$ matrix $\bar{F}_{t+1|t}$ as the corresponding unconditional period transition matrix as follows:*

$$\bar{F}_{t+1|t} =$$

$$
\begin{vmatrix}
\sum_{k_t=1}^{K} p(k_t|x_t^1)F^{k_t}_{t+1|t}(x^1_{t+1}|x^1_t) & . & \sum_{k_t=1}^{K} p(k_t|x_t^1,)F^{k_t}_{t+1|t}(x^{L_{t+1}}_{t+1}|x^1_t) \\
\vdots & & \vdots \\
\vdots & & \vdots \\
\sum_{k_t=1}^{K} p(k_t|x_t^{L_t})F^{k_t}_{t+1|t}(x^1_{t+1}|x^{L_t}_t) & . & \sum_{k_t=1}^{K} p(k_t|x_t^{L_t})F^{k_t}_{t+1|t}(x^{L_{t+1}}_{t+1}|x^{L_t}_t)
\end{vmatrix}
$$

For $t = 0, ...T-1$ define $\mathbf{F}_{t,t+s}$ as following:

$$
\begin{aligned}
\mathbf{F}_{t,t+r} &= \prod_{\tau=t}^{t+r-1} \bar{F}_{\tau+1|\tau} && \textit{for } r > 0 \\
\mathbf{F}_{t,t+r} &= I_{L_t} && \textit{for } r = 0
\end{aligned}
$$

where I_{L_t} is the identity matrix of size $L_t \times L_t$. The $F(x_s|x_t, k_{jt})$ given in (2.8) is equal to $\mathbf{F}_{t,s}$.

Definition 3. *Let $L_t \times 1$ vector $H_{k_t}(x_0'|x_T)$ is defined as $H_{k_t}(x_0'|x_T) = (H_{k_t}(x_0^1|x_T), ., H_{k_t}(x_0^L|x_T))'$ Define H^{k_T} as the $L_T \times L$ matrix of conditional transition matrix from the state in period T of current generation to period 0 of next generation given the choice k_T in period T. Finally let $H^{k_t}(x^m_0|x^l_T)$ denote the mth column and lth row of the matrix H^{k_T}.*

Definition 4. *Let the function $e_{k_t}(x_t, p(x_t))$ denote the expectation of the period unobservable (preference shock) conditional on x_t and k_t; $e_{k_t}(x_t, p(x_t)) = E[\varepsilon_{tk_t} \mid x_t, k_t]$. The vector $p(x_t)$ is the $(K-1)\times 1$ vector of conditional choice probabilities, $p(x_t) = (p(1|x_t), ..., p(K-1|x_t))'$ Let P_t be the $[L_t \times (K-1)] \times 1$ vector of conditional choice probabilities at period t. Let $\bar{u}_{tk_t}$, $\bar{p}_{tk_t}$ and $\bar{e}_{k_t}(P_t)$ be $L_t \times 1$ vectors that stack the corresponding elements at all states $((x_t) \in (x^1_t, ..., x^{L_t}_t))$ for alternative k_t in period t. Define the $L_T \times 1$ vector $\bar{n}_T = (n^1_T, ..., n^{L_T}_T)'$ as the corresponding number of children associated with all states in period T.*

Definition 5. *Let the $L \times L_T$ matrix $\mathbf{P}_{k_0,..k_T}$ be defined as follows:*

$$\mathbf{P}_{k_0,..k_T} = \bar{F}_{1|0}^{k_0} \otimes ...... \otimes \bar{F}_{T|T-1}^{k_{T-1}}$$

where we use the notation $\otimes$ to denote matrix multiplication whenever the content might not be clear enough to represent without it or there are both multiplications with scalars and matrices in the same expression. $*$ corresponds to element by element multiplication.

Definition 6. *The $L \times L$ intergeneration transition matrix $H_{k_0,..k_T}$ is defined as follows in terms of period primitives:*

$$H_{k_0,..k_T} = \left(\sum_{s=0}^{T} b(k_s)\right)^{1-v} \times \mathbf{P}_{k_0,..k_T} \otimes H^{k_T}$$

Definition 7. *The unconditional transition matrix $H^U(P)$ is defined as follows in terms of period primitives:*

$$H^U(P) = \sum_{\substack{k_0,..k_T \\ \in K \times ... \times K}} \bar{p}_{0k_0} * \mathbf{P}_{k_0,..k_T} \otimes \bar{p}_{T,k_T} * \left[\left(\sum_{s=0}^{T} b(k_s)\right)^{1-v} \times \mathbf{P}_{k_0,..k_T} \otimes H^{k_T} \right]$$

$H(x_0|x_t, k_{jt})$ given in (2.8) is just obtained by:

$$H(x_0|x_t, k_{jt}) = \sum_{k_{jt},..k_T} \mathbf{P}_{k_{jt},..k_T} \otimes \bar{p}_{T,k_T} * \left[\left(\sum_{s=t+1}^{T} [b(k_s) + b(k_{jt}) + N_t] \right)^{1-v} \times \mathbf{P}_{k_{jt},..k_T} \otimes H^{k_T} \right]$$

Replacing above definitions into the (A.1) leads to the representation given in (2.8).

The proof of the matrix equation given in (2.10) follows directly from the stationarity property and the representation which make use of the conditional choice probabilities in equation (2.9). For the derivation of this for infinite time single agent dynamic optimization problems, for instance see Aguirregabiria and Mira (2002)(1).

APPENDIX B

APPENDIX FOR CHAPTER 3

Data construction for the empirical application

In Appendix D, we describe in more detail the construction of our samples and the construction of the variables used in our study. We used data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID). We selected individuals from year 1968 to 1997 by setting the individual level variables "Relationship to Head" to head or wife or son or daughter. However we restricted the sample by dropping the individuals who are son or daughter and their ages are less than 17. We set "Sex of the Individual" to female and male and "Why Nonresponse" variable to zero categories which denotes individuals who were still members of a panel family in PSID. For the twenty-nine-year Family Individual Respondents File of the PSID, this initial selection produced a sample of 12,051 males and 12,744 females of which we have at least one year of observation in the years considered. The total number of observations we initially had for 24,795 individuals were 423,631. We lost observations due to missing data or inconsistent observations.

First type of missing data occurred with respect to the measure of the race. We used the family variable "Race of the Household Head" to measure the race variable in our study. There is a family variable that records information about the race of the wife separately, but this variable was included in the PSID only for the interviewing years 1985 and 1986. Defining the race variable in our empirical study as the race of the household head should not create much measurement error because the individuals in our sub-sample are either household heads themselves or wives of such heads. For the interviewing years 1968-1970, the values of 1 to 3 denote white, black, and Puerto Rican or Mexican, respectively. 7 denotes other (including Oriental and Philippino), and 9 denotes missing data. For 1971 and 1972, the third category is redefined as Spanish-American or Cuban, and between 1973-1984, just Spanish-American. After 1984, this variable was coded such that values of 1-4 correspond to the categories white, black, American Indian, Aleutian or Eskimo, and Asian or Pacific Islander, respectively. A value of 7 denotes the other category, and a value of 9 denotes missing data. Missing data in this variable results in a loss of 1,780 observations.

We used the individual level variables that indicate the educational attainment level of the household head or wife to measure the education variable. The variable "Completed Education" recorded in the individual part of the data record applies if the individual is a household head or wife. We used this individual level variable to construct the educational attainment of female. In order to minimize the observation losses due to this variable's applicability, we checked any other information in the family level education variables too. For both the head and wife, the coding of this variable is as follows: 1-16 highest grade or year of school completed, 17 at least some postgraduate work. A value of 90 denotes missing data. However, one difficulty in using the individual level education variable is that if the coding for this variable changes in year 1991 along with the codes for family level variables for education. Before 1991, for both the head and wife, the coding of this variable is as follows: 1: 0-5 grades, 2: 6-8 grades, 3: 9-11 grades, 4: 12 grades and no further training, 5: 12 grades plus nonacademic training, 6: College but no degree, 7: College BA but no advanced degree, and 8: College and advanced or professional degree. For both the head's and wife's education variable, a value of 9 denotes missing data. Therefore in order to obtain a consistent measure for this variable we did the following. First we found the most recent record for the individual in PSID. If this record is in year 1991 or beyond, we coded every education value in the previous years with this value if it is consistent across the years over 1991. If not we set the maximum of this variable as the completed education..We removed the 9,267 observations belonging to the individuals who have the final educational attainment variable as either 0 or greater than 30. The education is grouped into 4 discrete classes. We set the education be either LHS (less than high school) if education is less than 12 years; HS (high school) if education is 12 years; SC (some college) if education is more than 12 years but less than 16 years and COL (college) if education is equal or more than 16 years.

The marital status of a individual in our sub-sample was determined from the marital status of the head. This variable was coded differently for the interviewing year 1968, on the one hand, and the remaining years on the other. For 1968, the values 1 through 5 denote the categories married, single, widowed, divorced, and separated, respectively. The value 8 denotes married but spouse absent, and 9 denotes missing data. After 1968, the sixth category is dropped. There is 175 observations loss that can be attributed to this variable. The number of individuals in a household and the total number of children within that household were also determined from the family level variables of the same name. In 1968, a code for missing data (equal to 99) was allowed for the first variable, but in other years, missing data were assigned. The second variable, which indicates the total number of children under 18 in the family regardless of their relationship to the head,was truncated above at the value of 9 for the interviewing years 1968 and 1971. After 1975, this variable denotes the actual number of children within the family unit.

Labor income variables are constructed from the PSID variables "Total hours worked Head, "Hours Wife worked", "Labor income of Head", "Labor Income of Wife". The work hours reported in these variables are annual hours for the previous working year. Similarly the labor income is the reported labor income for the previous year. Therefore we matched those variables for the individual in a particular year by using the variable in the previous year. The labor earnings used in the analysis and part time and full time indicator variables are all obtained by these PSID variables. Total of 107,248 missing values generated due to missing observations and miscoded values and 946 miscoded values were deleted immediately but the remaining observations were kept with the missing values. The miscoded observations deleted were the ones reporting more than 4,380 hours annually for the total hours worked. The observations reporting positive earnings with a value of 0 for this variable are converted to missing. Also wage rate is calculated by dividing the annual labor earnings by the total annual hours. The observations for labor market variables were converted to missing if the wage rate was higher than $250/hour and lower than $1/hour. The part time variable is set equal to 1 if the individual worked less than 20 hours. week. Similarly the full time is constructed as equal to 1 if the weekly number of labor hours is more than or equal to 20.

Maternal time is proxied by the variables of the names;"Head Weekly Housework Hours" and "Wife Weekly Housework Hours" from the Family File of PSID. Those variables were missing in 1975 and 1982. For these years, they were interpolated from the available data for the previous and future values. Before 1993, the missing value for this variable was defined as 99, and after it were 112. Also before 1976 there were no record for this variable with the same name and definition in the Family file. However Individual File has an entry for "Weekly Housework". This variable has the same context with the above variable. We replaced the variable values with the ones from Individual File which were available for years 1969 to 1975. We had total of 102,364 missing and miscoded observations for the constructed variable. We deleted 6,951 of them but kept the rest with the missing values. The reason we kept the missing values for the maternal time and the labor income and participation variables is that estimation of earnings equation and the GMM estimation of the model's structural parameters (also the estimation of CCPs) requires different variables for estimation. For instance we don't need the maternal time for the estimation of the earnings equation; therefore we don't want to lose observations related to missing data due to maternal time. Obviously for some of the observations, both maternal time and labor income are missing, but we delay the deletion until the specific estimation. The discrete variable used in the structural estimation for the female's time investment is constructed from the maternal time investment variable by setting it equal to 0 if the number of weekly housework hours is less than 30 and setting it equal to 1 if it is equal or more than 30. The reason for choosing 30, instead of the labor market part time convention of 20 hours is due to the concentration of this variable around 20. Since in the estimation we want to reveal the effect of maternal investment, we wanted to separate the hours which is higher than the mean level of this variable.

For the first stage estimations, we apply further restrictions to the data in order to construct the variables needed for the different estimations. For instance the earnings equation requires the knowledge of past 4 participation decisions in the labor market. This immediately eliminates the individuals with less than 5 years of sequential observations in the data. We have a sample of 139,827 observations that fulfill all the data requirements for this estimation. The same condition applies to the base sample since the past four labor market part time and full time participation decisions are part of the state vector. However we do not need those variables for the individuals at the age of 17, since this is the starting age of the individual to the lifecycle in the empirical model and we assume the individual is not married and has not worked yet. Therefore this will enable us to use the observations of age 17 individuals with no labor past labor market outcomes. The argument follows similarly for age 18 individuals, since we only need to observe the age

17 labor market outcomes for them. This generates more data than the sample used in earnings equation estimation in this dimension, but now we need to keep track of the number of children and maternal time investment. In the application maternal hours should be observed as early as children of the female were born. Therefore we eliminated observations for the females who we observe at an age where they already have their children grown up. Also we deleted all sequence of observations for a female who has missing years in the sequence. For instance if we observe the female in years 1975, 1976, 1977, but then in years 1980, 1981,..., then we cannot include her in the analysis. Furthermore if we observe an individual, either male and female who is married, we need to have the certain spouse characteristics available in the data. These restrictions drop the sample to 136,916 observations. 56,812 observations are for 6,517 male and 80,104 observations are for 6,732 female individuals who fulfill all data requirements. This constitutes the base sample for the estimation of the structural parameters of the model.

Existence of Pure Strategy Equilibrium

Proof of Proposition 1. To show that the continuation values are super modular it suffices to show that the per-period utility is super modular and that the transition functions are super-modular. First we show that the per-period is super modular, i.e. $u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$ is super-modular in $k_{\sigma t}$ for any $x_{\sigma t}$ and $k_{-\sigma t}$ if;

$$u(k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + u(k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \geq u(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \text{ for all } (k'_{\sigma t}, k_{\sigma t}). \quad \text{(B.1)}$$

Without loss of generality let $k'_{\sigma t} \succeq k_{\sigma t}$, given that the choice set satisfies partial order

$$u(k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = u_{1\sigma t}(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + u_{2\sigma t}(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + \varepsilon_{k'_{\sigma t}} = u(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$$

and similarly

$$u(k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = u_{1\sigma t}(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + u_{2\sigma t}(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + \varepsilon_{k_{\sigma t}} = u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$$

Thus the condition holds.

Next we show that the transition functions are super-modular. Let $P_{Ft}(\widehat{X}|x,k)$ and $P_{MT}(\widehat{X}|x,k)$ be the probabilities of the set $\widehat{X} \subseteq X$ occurring with respect to $F(x_{t+1}|x_t, k_t)$ and $M(x'_0|x, D)$, i.e.

$$\begin{aligned} P_{Ft}(\widehat{X}|x,k) &= \sum_{x' \in \widehat{X}} F_t(x'|x,k) \\ P_{Ms}(\widehat{X}|x,k) &= \sum_{x'_0 \in \widehat{X}} M(x'_0|x,D) \end{aligned}$$

We say that $\widehat{X} \subseteq X$ is an increasing set if $x' \in \widehat{X}$ and $x'' \geq x'$ imply $x'' \in \widehat{X}$. Therefore $F_t(x'|x,k)$ and $M(x'_0|x,D)$ are stochastically super-modular in $k_{\sigma t}$ for any $x_{\sigma t}$ and $k_{-\sigma t}$ if:

$$\begin{aligned} & P_{Ft}(\widehat{X}|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + P_{Ft}(\widehat{X}|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \\ \geq \quad & P_{Ft}(\widehat{X}|k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + P_{Ft}(\widehat{X}|k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \text{ for all } (k'_{\sigma t}, k_{\sigma t}), \end{aligned} \quad \text{(B.2)}$$

and

$$\begin{aligned} & P_{Mt}(\widehat{X}|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + P_{Mt}(\widehat{X}|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \\ \geq \ & P_{Mt}(\widehat{X}|k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) + P_{Mt}(\widehat{X}|k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) \text{ for all } (k'_{\sigma t}, k_{\sigma t}) \end{aligned} \tag{B.3}$$

for any increasing set $\widehat{X} \subseteq X$. Without loss of generality assume that for $k'_{\sigma t} \geq k_{\sigma t}$, $F_t(x'|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_t) = F_t(x'|k'_{\sigma t}, k_{-\sigma t}, x_t)$ and $F_t(x'|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_t) = F_t(x'|k_{\sigma t}, k_{-\sigma t}, x_t)$, therefore

$$P_{Ft}(\widehat{X}|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = \sum_{x' \subseteq \widehat{X}} F_t(x'|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_t) = \sum_{x' \subseteq \widehat{X}} F_t(x'|k'_{\sigma t}, k_{-\sigma t}, x_t)$$

and

$$P_{Ft}(\widehat{X}|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = \sum_{x' \subseteq \widehat{X}} F_t(x'|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_t) = \sum_{x' \subseteq \widehat{X}} F_t(x'|k_{\sigma t}, k_{-\sigma t}, x_t)$$

and the condition is satisfied for. $M(x'_0|x, D)$ is defined in Equation 3.18. Recall that

$$\Pr(e'_\sigma \mid x_f, x_m, D_s(k'_{\sigma t} \vee k_{\sigma t}, x_t, k_{-\sigma t}) = \Pr(e'_\sigma \mid x_f, x_m, D_s(k'_{\sigma t}, x_t, k_{-\sigma t}))$$

and

$$\Pr(e'_\sigma \mid x_f, x_m, D_s(k'_{\sigma t} \wedge k_{\sigma t}, x_t, k_{-\sigma t}) = \Pr(e'_\sigma \mid x_f, x_m, D_s(k_{\sigma t}, x_t, k_{-\sigma t}))$$

Thus, $\Pr(e'_\sigma \mid x_f, x_m, D_\sigma)$ is stochastically super-modular in $k_{\sigma t}$ for any $x_{\sigma t}$ and $k_{-\sigma t}$. These conditions are trivially satisfied for $\Pr(\eta'_\sigma \mid e'_\sigma), \Pr(e'_{-\sigma 0} \mid e'_\sigma)$ from the conditional independence assumption. Therefore,

$$\begin{aligned} & M(x'_0|x, D_\sigma(k'_{\sigma t} \vee k_{\sigma t}, x_t, k_{-\sigma t})) = \Pr(e'_\sigma \mid x_f, x_m, D_s(k'_{\sigma t} \vee k_{\sigma t}, x_t, k_{-\sigma t}) \Pr(\eta'_\sigma \mid e'_\sigma) \Pr(e'_{-\sigma 0} \mid e'_\sigma) \\ = \ & \Pr(e'_\sigma \mid x_f, x_m, D'_s(k'_{\sigma t}, x_t, k_{-\sigma t})) \Pr(\eta'_\sigma \mid e'_\sigma), \Pr(e'_{-\sigma 0} \mid e'_\sigma) = M(x'_0|x, D_s(k'_{\sigma t}, x_t, k_{-\sigma t})) \end{aligned}$$

And similarly $M(x'_0|x, D_\sigma(k'_{\sigma t} \wedge k_{\sigma t}, x_t, k_{-\sigma t})) = M(x'_0|x, D_s(k_{\sigma t}, x_t, k_{-\sigma t}))$.Thus,

$$P_{Ft}(\widehat{X_0}|k'_{\sigma t} \vee k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = \sum_{x' \subseteq \widehat{X_0}} M(x'_0|x, D_\sigma(k'_{\sigma t} \vee k_{\sigma t}, x_t, k_{-\sigma t})) = \sum_{x' \subseteq \widehat{X}_0} M(x'_0|x, D'_s)$$

and similarly $P_{Ft}(\widehat{X_0}|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = P_{Ft}(\widehat{X}_0|k'_{\sigma t} \wedge k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$ for any set $\widehat{X}_0 \subseteq X$.

Next we need to show that condition **Condition (ID)** holds. For females, for any $k'_{ft} \succeq k_{ft}$, and given any $k'_{mt} \succeq k_{mt}, x_{ft}$ the continuation value $v(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$ has increasing differences for every state $x_{t,}$ and age $t \leq T$. First note that that the the per period utility $u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t})$ has increasing differences,

$$\begin{aligned} & u(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) - u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = \alpha_\sigma(w_{ft}(k'_{\sigma t}) - w_{ft}(k_{\sigma t})) + \alpha_{fN}(b_t(k'_{\sigma t}) - b_t(k_{\sigma t})) + \\ & \theta_{fk'_t} - \theta_{fk_t} + \varepsilon_{k'_{\sigma t}} - \varepsilon_{k_{\sigma t}} = u(k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) - u(k_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) \end{aligned}$$

Similarly for males for any $k'_{mt} \succeq k_{mt}$, and given any $k'_{ft} \succeq k_{ft}, x_{mt}$

$$u(k'_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) - u(k_{\sigma t}, k_{-\sigma t}, x_{\sigma t}) = \alpha_\sigma(w_{ft}(k'_{\sigma t}) - w_{ft}(k_{\sigma t})) +$$
$$\theta_{fk'_t} - \theta_{fk_t} + \varepsilon_{k'_{\sigma t}} - \varepsilon_{k_{\sigma t}} = u(k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma t}) - u(k_{\sigma t}, k'_{-\sigma t}, x_{\sigma t})$$

We begin by deriving that for period T, the conditions for increasing differences in $(k_{\sigma t}, k_{-\sigma t})$ of the continuation value. Note that it is also the per period utility, but unlike all other periods, it includes the expected valuations of the children.

$$\upsilon_\sigma(k'_{\sigma T}, k'_{-\sigma T}; x_T) - \upsilon_\sigma(k_{\sigma T}, k'_{-\sigma T}; x_T) = \left(u(k'_{\sigma T}, k'_{-\sigma T}, x_{\sigma T}) - u(k_{\sigma T}, k'_{-\sigma T}, x_{\sigma T})\right) +$$
$$\beta\lambda\left(\frac{(N_{\sigma T} + b'_T)^{1-v}}{(N_{\sigma T} + b'_T)}\overline{V}_{N\sigma}(k'_{\sigma t}, k'_{-\sigma t}; x_T) - \frac{(N_{\sigma T} + b_T)^{1-v}}{(N_{\sigma T} + b_T)}\overline{V}_{N\sigma}(k_{\sigma t}, k'_{-\sigma t}; x_T)\right)$$

We showed above that $u(k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma T}) - u(k_{\sigma t}, k'_{-\sigma t}, x_{\sigma T})$ exhibits increasing differences thus it is suffices to establishes conditions for the second element to exhibit increasing difference, that is that

$$\frac{(N_{\sigma T} + b'_T)^{1-v}}{(N_{\sigma T} + b'_T)}\overline{V}_{N\sigma}(k'_{\sigma T}, k'_{-\sigma T}; x_T) - \frac{(N_{\sigma T} + b_T)^{1-v}}{(N_{\sigma T} + b_T)}\overline{V}_{N\sigma}(k_{\sigma T}, k'_{-\sigma T}; x_T) \quad \geq$$
$$\frac{(N_{\sigma T} + b'_T)^{1-v}}{(N_{\sigma T} + b'_T)}\overline{V}_{N\sigma}(k'_{\sigma T}, k_{-\sigma T}; x_T) - \frac{(N_{\sigma T} + b_T)^{1-v}}{(N_{\sigma T} + b_T)}\overline{V}_{N\sigma}(k_{\sigma T}, k_{-\sigma T}; x_T)$$

First note that labor supply decisions only enter $u(k'_{\sigma t}, k'_{-\sigma t}, x_{\sigma T}) - u(k_{\sigma t}, k'_{-\sigma t}, x_{\sigma T})$, thus, we only need to verify the property for choices $(k'_{\sigma t} \geq k_{\sigma t})$ and $\left(k'_{-\sigma t} \geq k_{-\sigma t}\right)$ which have higher birth and time spent with children decisions. We begin with $(k'_{\sigma t} \geq k_{\sigma t})$ and $\left(k'_{-\sigma t} \geq k_{-\sigma t}\right)$ for which $k'_{\sigma t}, k'_{-\sigma t}$ have higher time spent with children (suppose birth decisions are similar). We need to show that

$$\left[\overline{V}_{N\sigma}(k'_{\sigma t}, k'_{-\sigma t}; x_T) - \overline{V}_{N\sigma}(k'_{\sigma t}, k_{-\sigma t}; x_T)\right] \geq \left[\overline{V}_{N\sigma}(k_{\sigma t}, k'_{-\sigma t}; x_T) - \overline{V}_{N\sigma}(k_{\sigma t}, k_{-\sigma t}; x_T)\right]$$

Note that $D_s(k_{\sigma t}, k_{-\sigma t})$, is increasing in $k_{\sigma t}, k_{-\sigma t}$. The above condition can be written as:

$$\sum_{s=0}^{T-1}\left[b_s \sum_\sigma I_{\sigma s} \sum_{x'_0} V_{\sigma s}(x'_0)\left(M(x'_0|x_T, D_s(k'_{\sigma T}, k'_{-\sigma T})) - M(x'_0|x_T, D_s(k_{\sigma T}, k'_{-\sigma T}))\right)\right] +$$
$$b_T \sum_\sigma p_\sigma \sum_{x'_0} V_{\sigma T}(x'_0)\left(M(x'_0|x_T, D_T(k'_{\sigma T}, k'_{-\sigma T})) - M(x'_0|x_T, D_T(k_{\sigma T}, k'_{-\sigma T}))\right) \geq$$
$$\sum_{s=0}^{T-1}\left[b_s \sum_\sigma I_{\sigma s} \sum_{x'_0} V_{\sigma s}(x'_0)\left(M(x'_0|x_T, D_s(k'_{\sigma T}, k_{-\sigma T})) - M(x'_0|x_T, D_s(k_{\sigma T}, k_{-\sigma T}))\right)\right] +$$
$$b_T \sum_\sigma p_\sigma \sum_{x'_0} V_{\sigma T}(x'_0)\left(M(x'_0|x_T, D_T(k'_{\sigma T}, k_{-\sigma T})) - M(x'_0|x_T, D_T(k_{\sigma T}, k_{-\sigma T}))\right)$$

Thus as long as $M(x'_0|x_f, x_m, D_s(k_{\sigma t}, k'_{-\sigma t})$ exhibits increasing differences in D, the condition is satisfied. Thus, as long as $V_{\sigma s}(x'_0)$ is weakly increasing in $\eta'_\sigma, Ed'_{-\sigma 0}, Ed'_\sigma$ and $\Pr(\eta'_\sigma \mid Ed'_\sigma)\Pr(Ed'_{-\sigma 0} \mid Ed'_\sigma)$ weakly increase in Ed'_σ the condition is that $\Pr(Ed'_\sigma \mid x_f, x_m, D_s)$ satisfied increasing differences which is satisfied by Assumption 1.. Therefore the valuation function is weakly increasing in x'_0.

Next consider $k'_{\sigma t} \geq k_{\sigma t}$ and $k'_{-\sigma t} \geq k_{-\sigma t}$ for which let $b'_T = 1$ and $b_T = 0$. We need to show that given the highest difference in time spent with kids in *one* period, the decline in the mean quality of any existing child is small enough. We already know that we have increasing differences for all other dimensions of the state space except for birth. Denote by $\underline{d}$ and $\overline{d}$ the lowest and highest investment level possible in one period by one spouse. Suppose spouse σ strategies $k'_{\sigma t} \geq k_{\sigma t}$ involve same d and only differ by birth decisions. Suppose $k'_{-\sigma t}$ involve $\overline{d}$ and that $k_{-\sigma t}$ involve $\underline{d}$, the condition needed for increasing differences is therefore

$$\begin{aligned} & \frac{(N_{\sigma T}+1)^{1-v}}{(N_{\sigma T}+1)}\left[\overline{V}_{N\sigma}(k'_{\sigma t},k'_{-\sigma t};x_T)-\overline{V}_{N\sigma}(k'_{\sigma t},k_{-\sigma t};x_T)\right] \\ \geq \ & \frac{(N_{\sigma T})^{1-v}}{N_{\sigma T}}\left[\overline{V}_{N\sigma}(k_{\sigma t},k'_{-\sigma t};x_T)-\overline{V}_{N\sigma}(k_{\sigma t},k_{-\sigma t};x_T)\right] \end{aligned}$$

Define the average quality of the stock of children:

$$\begin{aligned} \widehat{V}_{N_T}(k_{\sigma t},k'_{-\sigma t};x_T) \ \equiv \ & \frac{1}{N_{\sigma T}+1}\sum_{s=0}^{T-1}\left[b_s\sum_{\sigma}I_{\sigma s}\sum_{x'_0}V_{\sigma s}(x'_0)M(x'_0|x_f,x_m,D_s(\frac{d}{N_{\sigma T}},\frac{\overline{d}}{N_{\sigma T}}))\right] \\ & +\frac{1}{N_{\sigma T}+1}\sum_{\sigma}p_\sigma\sum_{x'_0}V_{\sigma T}(x'_0)M(x'_0|x_f,x_m,D_T(\frac{d}{N_{\sigma T}+1},\frac{\overline{d}}{N_{\sigma T}+1})) \end{aligned}$$

Then sufficient conditions for increasing differences are:

$$\begin{aligned} & \frac{(N_{\sigma T}+1)^{1-v}}{(N_{\sigma T}+1)}\left[(N_{\sigma T}+1)\left(\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}+1},\frac{\overline{d}}{N_{\sigma T}+1};x_T)-\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}+1},\frac{\underline{d}}{N_{\sigma T}+1};x_T)\right)\right] \ \geq \\ & \frac{(N_{\sigma T})^{1-v}}{N_{\sigma T}}\left[N_{\sigma T}\left(\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}},\frac{\overline{d}}{N_{\sigma T}};x_T)-\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}},\frac{\underline{d}}{N_{\sigma T}};x_T)\right)\right] \end{aligned}$$

Rearranging the condition for all $0 \leq N_{\sigma T} \leq T$:

$$\left(\frac{N_{\sigma T}+1}{N_{\sigma T}}\right)^{1-v} \geq \frac{\left(\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}},\frac{\overline{d}}{N_{\sigma T}};x_T)-\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}},\frac{\underline{d}}{N_{\sigma T}};x_T)\right)}{\left(\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}+1},\frac{\overline{d}}{N_{\sigma T}+1};x_T)-\widehat{V}_{N_T}(\frac{d}{N_{\sigma T}+1},\frac{\underline{d}}{N_{\sigma T}+1};x_T)\right)}$$

That is, the highest ratio of the right hand side is obtained for the largest difference in time investment of a spouse, for a one period investment, and a strategy of an individual in which the higher one has birth. The conditions says that the increase difference in average quality of a child cause be investment difference of $\frac{\overline{d}-\underline{d}}{N_{\sigma T}}$ versus $\frac{\overline{d}-\underline{d}}{N_{\sigma T}+1}$ is bounded by the left hand side (which takes the lowest value at $N_{\sigma T} = T$ by concavity assumption). Note that this assumption can be translated to an assumption on the transition function $M(x'_0|x_f,x_m,D_s(k'_{\sigma t},k'_{-\sigma t}) - M(x'_0|x_f,x_m,D_s(k_{\sigma t},k'_{-\sigma t})$. We already assumed that the marginal increase in investment in a child is weakly increasing in the existing stock of investment (and the spouse's investment), thus the left hand side of the above inequality is weakly larger than 1. The additional condition therefore bounds the increase in probability of outcomes as a function of a one period investment. In addition valuations functions of the child are weakly increasing in parental investment. Since consumption rises in wages and since education increase expected wage as well as spouses' education (assortative matching) and expected wage, this is satisfied.

Finally solving backwards, we established conditions for increasing differences of $\upsilon_\sigma(k'_{\sigma T}, k'_{-\sigma T}; x_T)$. Assuming that $F(x'_{t+1}|x_t, k_t)$ satisfies stochastic increasing differences, we show that for period $T-1$, the continuation value $\upsilon_\sigma(k_{\sigma T-1}, k_{-\sigma T-1}; x_{T-1})$ satisfies increasing differences in $(k_{\sigma T-1}, k_{-\sigma T-1})$. Thus, since $u(k'_{\sigma T-1}, k'_{-\sigma T-1}, x_{T-1})$ satisfies increasing differences, $F\left(x_T|x_{T-1}, k'_{T-1}\right)$ satisfies stochastic increasing differences and $\upsilon_\sigma(k'_{\sigma T}, k'_{-\sigma T}; x_T)$ also satisfies stochastic increasing differences, it is left to show that $p(k_T|x_T)$ in equation 3.5 satisfies stochastic increasing differences. Because $\varepsilon's$ are conditionally independent across spouses, time and choices, it suffices to show that the individual choice probabilities satisfy increasing differences:

$$p(k'_{\sigma T}|k'_{-\sigma T}, x_T) = \int \left[\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\upsilon_\sigma(k'_{\sigma T}, k'_{-\sigma T}; x_T) - \upsilon_\sigma(k_{\sigma T}, k'_{-\sigma T}; x_T) \geq \varepsilon_{\sigma k' t} - \varepsilon_{\sigma k' t}\} \right] dF_\varepsilon$$

That is

$$\sum_{k'_{\sigma T}} p(k'_{\sigma T}|k'_{-\sigma T}, x_T) - \sum_{k_{\sigma T}} p(k_{\sigma T}|k'_{-\sigma T}, x_T) \geq \sum_{k'_{\sigma T}} p(k'_{\sigma T}|k_{-\sigma T}, x_T) - \sum_{k_{\sigma T}} p(k_{\sigma T}|k_{-\sigma T}, x_T)$$

Define

$$\upsilon_\sigma(k'_{\sigma T}, k'_{-\sigma T}; x_T) - \upsilon_\sigma(k_{\sigma T}, k'_{-\sigma T}; x_T) \equiv \Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, x_T)$$

Thus, we need to show that

$$\begin{aligned}
&\int_\varepsilon \left[\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, x_T) \geq \varepsilon_{\sigma k' t} - \varepsilon_{\sigma k t}\} \right] dF_\varepsilon - \\
&\int_\varepsilon \left[\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k_{-\sigma T}, x_T) \geq \varepsilon_{\sigma k' t} - \varepsilon_{\sigma k t}\} \right] dF_\varepsilon = \\
&\int_\varepsilon \left[\prod_{k'_{\sigma T} \neq k_{\sigma T}} 1\{\Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, x_T) - \Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k_{-\sigma T}, x_T) \geq 0\} \right] dF_\varepsilon
\end{aligned}$$

Since for all $\left(k'_{\sigma T}, k'_{-\sigma T},\right) \geq \left(k_{\sigma T}, k_{-\sigma T},\right)$,

$$\Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k'_{-\sigma T}, x_T) - \Delta\upsilon_\sigma(k'_{\sigma T}, k_{\sigma T}; k_{-\sigma T}, x_T) \geq 0$$

And from conditional independence of $\varepsilon's$, $p(k'_{\sigma T}|k'_{-\sigma T}, x_T)$ has increasing differences. By backwards induction, the same proof applies for all $t < T-1$ thus the continuation value $\upsilon_\sigma(k_{\sigma T}, k'_{-\sigma T}; x_T)$ satisfies increasing differences for all $0 \leq t \leq T$.

By backwards induction, the same proof applies for all $t < T-1$ thus the continuation value $\upsilon_\sigma(k_{\sigma T}, k'_{-\sigma T}; x_T)$ satisfies increasing differences for all $0 \leq t \leq T$. $\square$